

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 24, 2005, 06:59:05 ; Search time 180 Seconds
(without alignments)

1300.112 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 NWIQNSILLIILIFASSIS.....VTPHCTSLSEIDALVNY 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2390	100.0	457	1	PGLR_LYCCH
2	1295.5	54.2	467	1	PGLR_LYCCH
3	1282	53.6	463	2	PGMR52
4	1204.5	50.4	444	2	PGMR52
5	1185.5	49.6	428	2	PGMR52
6	1185	49.6	438	2	PGMR52
7	1149	48.1	433	2	PGMR52
8	1136	47.5	462	1	PGLR_PERRA
9	1130	47.3	434	2	PGMR52
10	1128	47.2	461	2	PGMR52
11	1127	47.2	431	2	PGMR52
12	1115.5	46.7	426	2	PGMR52
13	1113	46.6	433	2	PGMR52
14	1108	46.4	458	2	PGMR52
15	1106	46.3	433	2	PGMR52
16	1103	46.2	433	2	PGMR52
17	1091	45.6	460	1	PGLR_MALDO
18	1073	44.9	460	2	PGMR52
19	1061.5	44.4	435	2	PGMR52
20	1059	44.3	199	2	PGMR52
21	1043	43.6	460	2	PGMR52
22	1034.5	43.3	405	2	PGMR52
23	1022.5	42.8	422	2	PGMR52
24	1014	42.4	429	2	PGMR52
25	998.5	41.8	200	2	PGMR52
26	988.5	41.4	444	2	PGMR52
27	925	38.7	452	2	PGMR52
28	914	38.2	444	2	PGMR52
29	898.5	37.6	459	2	PGMR52
30	840.5	35.2	468	2	PGMR52
31	825.5	34.5	435	2	PGMR52

32	807	33.8	481	2	PGMR52
33	802	33.6	426	2	PGMR52
34	796.5	33.3	514	1	PGMR52
35	794	33.2	408	2	PGMR52
36	789	33.0	514	2	PGMR52
37	788.5	33.0	514	2	PGMR52
38	788	33.0	457	2	PGMR52
39	787.5	32.9	434	2	PGMR52
40	787.5	32.9	514	1	PGMR52
41	787.5	32.9	514	2	PGMR52
42	787.5	32.0	507	1	PGMR52
43	749.5	31.4	161	2	PGMR52
44	741.5	31.0	491	2	PGMR52
45	736	30.8	423	2	PGMR52

ALIGNMENTS

RESULT 1	PGLR_LYCCH	STANDARD	PRT	457 AA.
AC	PGMR52			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Polygalacturonase 2A precursor (EC 3.2.1.15) (PG-2A) (Pectinase).			
OS	Lycopersicon esculentum (Tomato).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; asterids;			
OC	Lamiidae; Solanales; Solanaceae; Solanum.			
OX	NCBI_TaxID=4081;			
RN	(1)			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=cv. Alisa Craig;			
RA	Bird C.R., Smith C.J.S., Ray J.A., Mourreau P., Bevan M.W., Bird A.S.,			
RA	Hughes S., Morris P.C., Grierson D., Schuch W.;			
RT	"The tomato polygalacturonase gene and ripening-specific expression in			
RT	transgenic plants.";			
RL	Plant Mol. Biol. 11:651-662(1988).			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Alisa Craig;			
RX	MEDLINE=87066731; PubMed=3786135;			
RA	Grierson D., Tucker G.A., Keen J., Ray J., Bird C.R., Schuch W.;			
RT	"sequencing and identification of a cDNA clone for tomato			
RT	polygalacturonase.";			
RL	Nucleic Acids Res. 14:8595-8603(1986).			
RN	(3)			
RP	SEQUENCE FROM N.A.			
RA	Sheehy R.E., Pearson J., Brady C.J., Hatt W.R.;			
RT	"Molecular characterization of tomato fruit polygalacturonase.";			
RL	Mol. Gen. Genet. 208:30-36(1987).			
RN	(4)			
RP	REVISIONS.			
RA	Hatt W.R.;			
RT	Submitted (OCT-1987) to the EMBL/GenBank/DBJ databases.			
RT	"FUNCTION: Acts in concert with the pectinesterase, in the ripening			
RT	process. Is involved in cell wall metabolism, specifically in			
RT	pectinolytic degradation.			
RT	CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-			
RT	galactosiduronic linkages in pectate and other galacturonans.			
RT	SUBCELLULAR LOCATION: Secreted.			
RT	DEVELOPMENTAL STAGE: In ripening fruit.			
RT	BIOTECHNOLOGY: The effect of PG can be neutralized by introducing			
RT	an antisense PG gene by genetic manipulation. The Flavr Savr			
RT	tomato produced by Calgene (Monsanto) in such a manner has a			
RT	longer shelf life due to delayed ripening.			
RT	- SIMILARITY: Belongs to the glycosyl hydrolase 28 family.			
RT	- SIMILARITY: Contains 4 PstI repeats.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL; X14074; CA32235.1; -
 DR EMBL; M37304; AAA34178.1; -
 DR EMBL; X04583; CA28254.1; -
 DR EMBL; X05656; CA29148.1; -
 DR EMBL; A15981; CA01256.1; -
 DR EMBL; A24194; CA01720.1; -
 DR PIR; A25534; A25534; -
 DR InterPro; IPR000743; Glyco_hydro_28.
 DR InterPro; IPR006626; PBH1.
 DR InterPro; IPR011050; Pectin_lyase_like.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR SMART; SM00710; PBH1; 4.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 DR Cell wall; Fruit ripening; Genetically modified food; Glycoprotein;
 KM Glycosidase; Hydrolase; Repeat; Signal.
 FT SIGNAL 1 24 Potential.
 FT PROPEP 25 71
 FT CHAIN 72 457 Polygalacturonase 2A.
 FT REPEAT 228 255 PBH1 1.
 FT REPEAT 256 277 PBH1 2.
 FT REPEAT 309 330 PBH1 3.
 FT REPEAT 338 359 PBH1 4.
 FT ACT SITE 293 293 Probable.
 FT CARBOHYD 189 189 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 240 240 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 286 286 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 311 311 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 457 AA; 50051 MW; 449E4DC369198074 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 2390; DB 1; Length 457;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYIQNSILLIIIFASSISTCRSNVIDNLFKQVYNNIIIEGFADHFOAYLSYSKNT 60
 DB 1 MYIQNSILLIIIFASSISTCRSNVIDNLFKQVYNNIIIEGFADHFOAYLSYSKNT 60
 QY 61 SNNNIDKVDKNGIVINVLSPGAGDGTNDNIAFEQAMNEACSSRTPVQFVVKKNY 120
 DB 61 SNNNIDKVDKNGIVINVLSPGAGDGTNDNIAFEQAMNEACSSRTPVQFVVKKNY 120
 QY 121 LKQITFGSPCRSSISVKIFGSLKSSKISDYKDRRLMIADSVONLVGGGTINGNGV 180
 DB 121 LKQITFGSPCRSSISVKIFGSLKSSKISDYKDRRLMIADSVONLVGGGTINGNGV 180
 QY 181 WMPSSCKINKSLPERDAFTALTWNCKNLKVNNLKSKNAQOIHKIPESCTNVVASNLMIN 240
 DB 181 WMPSSCKINKSLPERDAFTALTWNCKNLKVNNLKSKNAQOIHKIPESCTNVVASNLMIN 240
 QY 241 ASAKSPNTDGVHVSNTQYIOISDTIIGTDDDCISYSSGQNVQANTTCGPHGISISL 300
 DB 241 ASAKSPNTDGVHVSNTQYIOISDTIIGTDDDCISYSSGQNVQANTTCGPHGISISL 300
 QY 301 GSNSNSAYVSNVTNVEAKIIGAENGVRITKMGSGGQASNIKELVNEQDVKPIIIDON 360
 DB 301 GSNSNSAYVSNVTNVEAKIIGAENGVRITKMGSGGQASNIKELVNEQDVKPIIIDON 360
 QY 361 YCDRVPCIQOFSAVOVKVYVNTKTSATVAKKPCDCTNPFCEGIMENINLVGESSG 420
 DB 361 YCDRVPCIQOFSAVOVKVYVNTKTSATVAKKPCDCTNPFCEGIMENINLVGESSG 420
 QY 421 KRSSEATCKNVHNNAAEHVTPHCTSLSEISDEALLYNY 457
 DB 421 KRSSEATCKNVHNNAAEHVTPHCTSLSEISDEALLYNY 457

RESULT 2

PGLR ACTCH
 ID PGLR ACTCH STANDARD; PRT; 467 AA.
 AC P35316;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase).
 OS Actinidia chinensis (Kiwi) (Yangtze).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Ericales; Actinidiaceae; Actinidia.
 OX NCBI_TaxID=3625;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Deliciosa;
 RX MEDLINE=94302157; Pubmed=8029342; DOI=10.1104/pp.103.2.669;
 RA Atkinson R.G., Gardner R.C.;
 RA "A polygalacturonase gene from kiwifruit (Actinidia deliciosa).";
 RL Plant Physiol. 103:669-670(1993).
 CC -1- FUNCTION: Acts in concert with the pectinesterase, in the ripening process. Is involved in cell wall metabolism, specifically in polyuronide degradation.
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-galactosiduronic linkages in pectate and other galacturonans.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DEVELOPMENTAL STAGE: In ripening fruit.
 CC -1- SIMILARITY: Belongs to the glycoyl hydrolase 28 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; L12019; AAC14453.1; -
 DR InterPro; IPR000743; Glyco_hydro_28.
 DR InterPro; IPR006626; PBH1.
 DR InterPro; IPR011050; Pectin_lyase_like.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR SMART; SM00710; PBH1; 4.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 DR Cell wall; Fruit ripening; Glycoprotein; Glycosidase; Hydrolase;
 KM Signal.
 FT SIGNAL 1 27 Potential.
 FT CHAIN 28 467 Polygalacturonase.
 FT ACT SITE 306 306 Probable.
 FT CARBOHYD 290 290 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 467 AA; 50776 MW; 5A9A61483C028B7A CRC64;

Query Match
 Best Local Similarity 54.2%; Score 1295.5; DB 1; Length 467;
 Matches 257; Conservative 75; Mismatches 112; Indels 27; Gaps 7;

QY 1 MYIQNR-SILLIIIFASSISTCRSNVIDNLFKQVYNNIIIEGFADHFOAYLSYSKNT 58
 DB 1 MALQRRFQVITLIIIFSSIFIGYTSVAHED---PPHDYLF-EYGVDFAYSYITTI 55
 QY 59 IESNNNIDVNDNGI-----KYINVLSPFAKDDGTYNDIAFEQAMNE 101
 DB 56 GNDPFGSSSHENGIFGLRKVDYGMKRVLDASKTVNVDFFAKDKGD-DTKAFKAMKA 114
 QY 102 ACSSRTPVQFVVKKNYLLKQITFGSPCRSSISVKIFGSLKSSKISDY-KDRRLMIAD 160
 DB 115 ACSSTSSAVLLVLPK-KNVLVPRISFGSPCKSLMGWYGTITKASDDRSVDRKGRHVLV 173
 QY 161 DSVQNLVVGSGGTINGNGVWMPSSCKINKSLPERDAFTALTWNCKNLKVNNLKSKNAQ 220
 DB 174 DSVQNLVREGSGGTINGNGKIMQNSCKTNKALPCKAPFTALTFRYSKRVIVKMLKTENAQ 223
 QY 221 QIHKIPESCTNVVASNLMINASAKSPNTDGVHVSNTQYIOISDTIIGTDDDCISYSSGQ 280

Db	234	QIHVSPDNCVNWQASNLMTWATAPENSPNTDGHVHTGQIWHIISCVIGIHGDUCISIVNGSR	293
Qy	281	NVOATNITTCGPGHGISIGSLGSGSEAYVSNVTNEAKITAEENGVRITKYQGGSGGASN	340
	:		:
	:		:
Db	294	KRVANDITCGPGHGISIGSLGYGSEBAHVSDDVWNGAKLCTGTGTVIRIKTWQGGSGGASN	353
Qy	341	IKELANEMODVKKPIIIDQNYCDRVEPCIQPSANQVKNVYENIKGSAKVAIKRDCS	400
	:		:
	:		:
Db	354	IKQNVEMHNVEPIIIDQNYCDQKPCQEOSSAQVKNVYQNIKGTCSANVAITPDCS	413
Qy	401	TNPPCEGIIMENINLVGSEKPSBAKCKVHNNAEHTTPHCTSLSEIASEDE	451
	:		:
	:		:
Db	414	KRPCCGIVLEVDLEIEGGAARAKLCCNNVELSETGVVSPHCQEGGGBEEF	464

[illegible]

Db	174	DSQMLRVEGGGTTINGNQIMWMSCKTNKTLPECKDAPTALTFYKSGHVIWVKLAKTENAQ	233
Qy	221	QHIIKESTQTVVANKMLIMNASAKSPNMDGCVHVSNTQYIQISDPIIGTSGDCCSYVSGQ	280
Db	234	QIHVSFDNCVNWQASNLNWTTPENSPNDGJHVTGQIMHISCVIETSGDCCSYVSGQ	293
Qy	281	NYQATNTTCPGHGHSIGSLSGNSGEAVSVNVTNEAKIIGAENVIKTMOGSGGQASN	340
Db	294	KRVNDITGCGHGISISLGVGNSEAHVDVVVYNGAKLGGTTGVARIKTMOGSGGASN	353
Qy	341	IKFLVEMODVKYPIIIDQNTCDRVEPTIQPSAVQVKNVYENIKTSATKVALKDCS	400
Db	354	IKFQNVEMHNVNPIIIDQNYCDDKPCQEGSSAVQVKNIFYKNIKTCASNAVITFDCS	413
Qy	401	TMPGCGIIMENIINVGSGKPSKESTGCVHFNNAHVTPHCTSLSEIDEA	452
Db	414	KFPFCGIVLEVDLEITGGAAAKKLCNNVLSLSTGVSHPCSP--GESEA	462

RESULT 4	ID	Q94B15	PRELIMINARY;	PR1;	444 AA.
DT	AC	Q94B15			
DT	01-DEC-2001	(TREMBLrel)	19,	Created)	
DT	01-DEC-2001	(TREMBLrel)	19,	Last sequence update)	
DT	01-MAR-2004	(TREMBLrel)	26,	Last annotation update)	
DE		Polygalacturonase PGL			
OS		Vitis vinifera (Grape)			
OC		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC		Vitaceae; Vitis.			
OX		NCBI_TaxID=29760;			
RP		SEQUENCE FROM N.A.			
RX		MEDLINE=21658619; PubMed=11800390;			
RX		Nunan K.J., Davies C., Robinson S.P., Fincher G.B.;			
RT		"Expression patterns of cell wall-modifying enzymes during grape berry			
RL		development ";			
RN		Planta 214:257-264(2001).			
RN		[2]			
RP		SEQUENCE FROM N.A.			
RA		Nunan K.J., Davies C., Robinson S.P.;			
RL		Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
CC		- - SIMILARITY: Belongs to family 28 of glycosyl hydrolases.			
DR		EMBL; AY043233; AAK81876.1; -.			
DR		GO; GO:0005618; C:cell wall; IEA.			
DR		GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.			
DR		GO; GO:0004650; F:polygalacturonase activity; IEA.			
DR		GO; GO:0005975; P:carbohydrate metabolism; IEA.			
DR		InterPro; IPR000743; Glyco_hydro_28.			
DR		InterPro; IPR006626; PBH1.			
DR		InterPro; IPR011050; Pectin_lyas_like.			
DR		InterPro; IPR00408; Reg_chf condens.			
DR		Pfam; PF00295; Glyco_hydro_28; 1.			
DR		SMART; SM00710; PBH1; 4.			
DR		PROSITE; PS00502; POLYGALACTURONASE; 1.			
DR		PROSITE; PS00626; RCCL_2; UNKNOWN_1.			
KW		Cell wall; Glycosidase; Hydrolase.			
SQ		SEQUENCE 444 AA; 47419 MW; D94259F4792EF132 CRC64;			
Query Match		50.4%; Score 1204.5; DB 2; Length 444;			
Best Local Similarity		61.5%; Pred. No. 4.56-76;			
Matches 228; Conservative		53; Mismatches 87; Indels 3; Gaps 3			
QY		73 IKYINVLVSFGKGGGKTYNDIAPEQANNEACSSRTPVQVFVPPKKNKNVLLKQITFGSPCKS			132
DB		75 VKKNVNVYGGKGGDATE-AFKGAKKACSSPGSV-LVPPKKNVLLKQITFGSPCKS			132
QY		133 SIATVYFGSLEASSKISDY-KDRRLMTAFDSQVNLVVGGGGTINGNGVWMPSSCKINKS			191
DB		133 SIATVYIGTVQASTDRSATSNDMTHWLLFENVQNLAVGGGTINGNGTWTWENSCKAVYD			192
QY		192 LPCHDAFTALTFFNNCKNLKVNNNLKSRAAQDIIHFESCTNVVASNMIMNAAKSPNTGCV			251

```

Db      193  LPCGAPFALTFFYCKNLAVKLNKIQNAQOMHSEFKCVGASGLTITAPGNSPNTDGI 252
Qy      252  HVSNTQYIOISDPTTIGTDDCISIVSGSQNVQATNTTCGPGHGISIGSISGNSRAYVN 311
Db      253  HVSNTKTIQISSSVTIGDDCISIVSGSQNVQATNTTCGPGHGISIGSISGNSRAYVN 312
Qy      312  VTYNEAKTIGAEVGRKTKWGGSGQASNIKFLNTEMODVYKPTIITDNYCDRVEPCTIQ 371
Db      313  VTYNGVITLSTGTGTVKIKTMWGGSSASNIKFLNTEMODVYKPTIITDNYCDRVEPCTIQ 372
Qy      372  FSAVQKVVYENIKTSATKVAIKFDCSTNPFCEGIIMENINLVGSGKPSBATCQNVH 431
Db      373  SSAVQVQNVLYQNIKTSSSKKALISLDCSAKFPCCGILLRIDIKVGGKAAKAVCSNAR 432
Qy      432  FNNABHTVPHC 442
Db      433  VTMGVDSVENC 443

```

RESULT 5

```

ID      0944B5  PRELIMINARY;  PRT;  428 AA.
AC      0944B5;
DT      01-DEC-2001 (TREMBLrel. 19, Created)
DT      01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT      01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE      Dehiscence-related endopolysaccharuronase.
GN      Name=SDPG;
OS      Glycine max (Soybean).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX      NCBI_TaxID=3847;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Christiansen L.C., Dai Degang F., Ulvskov P., Borkhardt B.;
RT      "Examination of the dehiscence zone in soybean pods and isolation of a
RT      dehiscence-related endopolysaccharuronase gene.";
RL      Plant Cell Environ. 25:479-490(2002).
CC      -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
DR      EMBL, AF434714; AAL30418.1; -.
DR      GO; GO:0005618; C:cell wall; IEA.
DR      GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR      GO; GO:0004650; F:polysaccharuronase activity; IEA.
DR      GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR      InterPro: IPR000743; Glyco_hydro_28.
DR      InterPro: IPR006626; PBH1.
DR      InterPro: IPR011050; Pectin_lyas_like.
DR      Pfam; PF00295; Glyco_hydro_28; 1.
DR      SMART; SM00710; PBH1; 5.
DR      PROSITE; PS00502; POLYGALACTURONASE; 1.
KW      Cell wall; Glycosidase; Hydrolase.
SQ      SEQUENCE 428 AA; 45830 MW; F553DBA6C297626D CRC64;

```

Query Match
Best Local Similarity 49.6%; Score 1185.5; DB 2; Length 428;
Matches 229; Conservative 79; Mismatches 116; Indels 21; Gaps 6;

```

Qy      1  MVIDGRNIIILTIIFASSISTCRSNVTDNLFKQVYDNLIEQEFADHPQAYLSYLSKITE 60
Db      1  MALQRRHLISLTFIMF--SFVACHSTVLDDG-----GCTFKDPIKQKSDTVM 44
Qy      61  SNNNIDKVD--KNGIRIVNLSFGAKGDKGTVDNTAFQANMEACSSRTTPQVFPKKNYL 118
Db      45  STRKFKLGDISSLKTVVNDYGAAPGQKTDIDTAFKAEVACSSGGA-VFVPR-KN 102
Qy      119  YLLKQITFGSPCRSSISVYIFGSLASKISDY-KDRRLWTFADSVQNLVVGSGGTINGN 177
Db      103  YLLKPTFGSPCESDIEVQISGILEASNLDSYSEDLTHMLVPSIKLSVKGSGGTIDGN 162
Qy      178  GQVWPPSCSKINKSLPCRDAFTALTFFWNCNKLKVNNLKSKNAQOIHKFESCTNVVANSNL 237

```

```

Db      163  GNIMWONSCKVNEKLPCKNAFTALTFFYCKDULVEDLITKNGQOMOVFQNSNVQVSDL 222
Qy      238  MINAAGSPNTDGVVNSVTOYIOISDPTTIGTDDCISIVSGSQNVQATNTTCGPGHGISI 297
Db      223  TTTAPGDEPNPDGHTVNTQNIQISNSVITGTDDCISIVSGSKVDVLTADIIICPGHGISI 282
Qy      298  GSLGSGSEAVSVNTVNEAKTIGAEVGRKTKWGGSGQASNIKFLNTEMODVYKPTI 357
Db      283  GSLGSGSKPVSQITVYKGAQLSGTTGKIKTKWGGSSASNIQFQNTQMDNVANPIT 342
Qy      358  DQNYCDRVEPCTIQFSAVQKVVYENIKTSATKVAIKFDCSTNPFCEGIIMENINLVG 417
Db      343  DQNYCDDETPEBRTSAVQKRVNLYQNIKTSSSKKALISLDCSAKFPCCGILLRIDIKL 402
Qy      418  ESKGPEATCKNVHFNNAHTVPHC 442
Db      403  EGGGEAKASCNISVELSYRGDVNPLC 427

```

RESULT 6

```

ID      09SEB7  PRELIMINARY;  PRT;  438 AA.
AC      09SEB7;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE      Putative polysaccharuronase.
GN      Name=F17A17.31;
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsie.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Lin X., Kaul S., Town C.D., Benito M.-I., Craasy T.H., Haas B.,
RA      Rohnig C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA      Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL      Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC      -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
DR      EMBL, AC013483; AF21207.1; -.
DR      GO; GO:0005618; C:cell wall; IEA.
DR      GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR      GO; GO:0004650; F:polysaccharuronase activity; IEA.
DR      GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR      InterPro: IPR000743; Glyco_hydro_28.
DR      InterPro: IPR006626; PBH1.
DR      InterPro: IPR011050; Pectin_lyas_like.
DR      InterPro: IPR000408; Reg_chit_condens.
DR      Pfam; PF00295; Glyco_hydro_28; 1.
DR      SMART; SM00710; PBH1; 4.
DR      PROSITE; PS00502; POLYGALACTURONASE; 1.
DR      PROSITE; PS00626; RCCL_2; UNKNOWN_1.
KW      Cell wall; Glycosidase; Hydrolase.
SQ      SEQUENCE 438 AA; 48444 MW; ACAAG960403EB37 CRC64;

```

Query Match
Best Local Similarity 49.6%; Score 1185; DB 2; Length 438;
Matches 230; Conservative 76; Mismatches 121; Indels 16; Gaps 5;

```

Qy      8  ILLTIIFASSISTCRSNV--IDNLFQK--VYDNLIEQEFADHPQAYLSYLSKITES 61
Db      5  ILLISVPLTLTFPSGCFSSYFNNHRDDLFWSSNVLYETRNQHGHN-----TRNSHL 56
Qy      62  NNNIDKVDKNGIKIVNLSFGAKGDKGTVDNTAFQANMEACSSRTTPQVFPKKNYL 121
Db      57  KNRGVAPRSPSPFNVTGAKANND--DSKAFKAMAEACSSGTGIYIYAPKRRDYL 115
Qy      122  KQITFGSPCRSSISVYIFGSLASKISDYKDRRLWTFADSVQNLVVGSGGTINGNGQV 181
Db      116  KAVTFSGPCSSSLITLYGRIEAWENPDSYKERRHRIWFENVNVLRVVGGGRIDQNGIHW 175
Qy      182  WPSSCKINKSLPCRDAFTALTFFWNCNKLKVNNLKSKNAQOIHKFESCTNVVANSNLMA 241

```

```

Db 176 WPKSCKINPOLPCIGAPTAVTFVECNRLRVSNIRLENAQOMHLTFQCKKVKALNLMVTS 235
Qy 242 SAKSPNTDGHVANTQYIOISDTIIGTDCICISVSSQNVQATNTICGPHGISIG 301
Db 236 PASPNTDGHVANTQYIOISDTIIGTDCICISVSSQNVQATNTICGPHGISIG 295
Qy 302 SGSEAVSVNTVEAKIIGAENGVRKKTQGGSGGASNIKFLNVEMQDYKYPPIIDONY 361
Db 296 EDSSEAVSVNTVEAKIIGAENGVRKKTQGGSGGASNIKFLNVEMQDYKYPPIIDONY 355
Qy 362 CDRVEPCIQFSAVOQNVVYENIKGTSAATKVAIKPDCSTNPFCEGIMENINLVGE 421
Db 356 CDRVEPCIQFSAVOQNVVYENIKGTSAATKVAIKPDCSTNPFCEGIMENINLVGE 415
Qy 422 P-SEATCKNVHFNNAEHVPCHT 443
Db 416 DVSKASCKNVKLDTRGVNSPLCT 438

RESULT 7
Q8RY29 PRELIMINARY; PRT; 433 AA.
ID 08RY29
AC 08RY29;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE AC2941850/111A7.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shim P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh U.,
RA Bower L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Becker J.R.;
RL Submitted (EBE-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC EMBL: AY078936; AL84942.1; -.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO: GO:0004650; F:polysaccharonase activity; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR000743; Glyco_hydro_28.
DR InterPro: IPR006626; Pbh1.
DR InterPro: IPR011050; Pectin_lyase_like.
DR InterPro: IPR000408; Reg_chf_condens.
DR Pfam: PF00295; Glyco_hydro_28; 1.
DR SMART: SM00710; Pbh1; 5.
DR PROSITE: PS00502; POLYGALACTURONASE; 1.
DR PROSITE: PS00628; RCT1_2; UNKNOWN_1.
DR Cell wall; Glycosidase; Hydrolase.
SQ SEQUENCE 433 AA; 46492 MW; 08E1EE62AA05A26B CRC64;

Query Match 48.1%; Score 1149; DB 2; Length 433;
Best Local Similarity 50.7%; Pred. No. 3.3e-72;
Matches 226; Conservative 76; Mismatches 118; Indels 26; Gaps 6;

```

```

Qy 121 LKQITPSPCRSSISYKIFGSLSEASGISDYKDRRLMAFDVQNLVWGG--GTINGNG 178
Db 113 LKSIQLTGPNSTLTVQIFGTLISAQKRSYKDISKIMFDPGVNLSVDGGDTGVVDGNG 172
Qy 179 QVWPPSSCKINSLPRDAPALTATFNWCKLKNLKSXAQOIHKFEECTVWASNL 238
Db 173 ETWQNSCKRNKAPCTXAFTALTFFNSKSLYKNLKVRAQOIQISIEKSNVQVSNV 232
Qy 239 INSAKSPNTDGHVANTQYIOISDTIIGTDCICISVSSQNVQATNTICGPHGISIG 298
Db 233 VTPASPNTDGHVANTQYIOISDTIIGTDCICISVSSQNVQATNTICGPHGISIG 292
Qy 299 SLGSGSEAVSVNTVEAKIIGAENGVRKKTQGGSGGASNIKFLNVEMQDYKYPPIID 358
Db 293 SLGDDNSKAFVSGVTVTDGAKLSTGDNVRIKTYQGGSGATSNIFQNIQMDVNVKPIID 352
Qy 359 QNYCDRVEPCIQFSAVOQNVVYENIKGTSAATKVAIKPDCSTNPFCEGIMENINLVGE 418
Db 353 QDDCDKQK-CTTEKSAVOQNVVYRDISGTSASENAITFNCSTKVYPCQGIIVLDVRIKQ- 410

Qy 419 SGKPSATCKNVHFNNAEHVPCHT 444
Db 411 ---GRATCTNANVVDKGAVLPCNS 432

```

```

RESULT 8
PGRL_PEAR STANDARD; PRT; 462 AA.
ID 002096;
AC 002096;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE POLYGALACTURONASE precursor (EC 3.2.1.15) (PG) (Pectinase).
OS Persea americana (Avocado).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; magnoliids; Laurales; Lauraceae; Persea.
OX NCBI_TaxID=435;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Hass; TISSUE=pericarp;
RX MEDLINE=93184201; PubMed=8095163;
RA Doplo B., Lowe A.L., Wilson I.D., Merodio C., Grierson D.;
RT "Cloning and characterization of avocado fruit mRNAs and their
RT expression during ripening and low-temperature storage.";
RL Plant Mol. Biol. 21:437-449(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Hass; TISSUE=Mesocarp; DOI=10.1104/pp.103.1.289;
RX MEDLINE=94269193; PubMed=8208850;
RA Kutsunai S.Y., Lin A.C., Percival F.W., Latties G.G.,
RA Christoffersen R.E.;
RT "Ripening-related polygalacturonase cDNA from avocado.";
RL Plant Physiol. 103:289-290(1993).
CC -1- FUNCTION: Acts in concert with the pectinesterase, in the ripening
CC process. Is involved in cell wall metabolism, specifically in
CC polyuronide degradation.
CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galacturonidic linkages in pectate and other galacturonans.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: In ripening fruit.
CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 28 family.
CC -1- SIMILARITY: Contains 5 Pbh1 repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66426; CAA47055.1; -.
DR EMBL: L06094; AAA32914.1; -.

```

DR EMBL: AJ250919; CAC05658.1; -.
DR EMBL: AJ250919; CAC05657.1; -.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.
DR GO: GO:0004650; F:polygalacturonase activity; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPRO0743; Glyco_hydro_28.
DR InterPro: IPRO06626; Pbh1.
DR InterPro: IPRO11050; Pectin_1yaa_like.
DR InterPro: IPRO00408; Reg_chit_condens.
DR Pfam: PF00295; Glyco_hydro_28; 1.
DR SMART: SM00710; Pbh1; 5.
DR PROSITE: PS00502; POLYGLACTURONASE; 1.
DR PROSITE: PS00626; RC01_2; UNKNOWN_1.
DR Cell wall; Glycosidase; Hydrolyase.
KW SEQUENCE 434 AA; 46603 MW; E9249AC1BEC0C219 CRG4;

Query Match	47.3%;	Score 1130;	DB 2;	Length 434;
Best Local Similarity	49.8%;	Pred. No. 7.1e-71;		
Matches 221; Conservative	78;	Mismatches 117;	Indels 28;	Gaps 7

```

QY      8 ILLIIIIIFASSISTCNSNVIDNLFKQYVNNILBEOFAHDFQAYLS-YLSKNIESSNNID 66
Db      11 LTVMSIFMLTSMCEASRNNI-----NRDHSYGTFFSNLLIGRPDIITLK 55
QY      67 KYDKNGIKV---INVLSPGAKGDGKTYNIAPEQAMNAGCSRPTEQFVVRKNNYLLKQ 123
Db      56 SVARASLRLEPTVSVDPGAKDGKTDITQAFVNMKAKCSSGSAVNLVBEGETYFLKS 115
QY      124 IFFSGPCCSSISVKIFGSLBASSKISDYKDRRLIAEDSVONLVAVGG--GTINGNG-QV 180
Db      116 LRNPFCSSVTLVQLGLTFLSAGQSRSDYEDLSKITFDGVNSLTVDEGAGGTGVNGNAET 175
QY      181 WPPSSCKINKSLPCRDAPALTATFPNCKNKLKNNLSKNAQOOIHKPESCINVAASNIMIN 240
Db      176 WMESSCKRNEAKCKTCAPTALTFPNSNKLKNNLRVYDAQOIOISLEKCSNVQSVNEVET 235
QY      241 ASASPTMDGVHVSNTQYIOISTIIIGTDGDISIVSGSONVQATNITCGPHGISISGL 300
Db      236 APASPTMDGHTHTTONTQIVNSNIIIGTDGDISISSGSONQVINDLTGPHGISISGL 295
QY      301 GSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGQANINKEPLANEMODVKYPIIIID 360
Db      296 GDNNSKAFVSGVTVDGAKLSTGDNVRITKMOGSGGTASNIIIFONIQMENVNPIIIID 355
QY      361 YCDRVEPCIOFSAVOYKNNVVENIKGISATYALKPDCSTNPFCEGLIMENINLVGEG 420
Db      356 YCDK-SCITEQSAVOIKNNVVRNISIGTSASIDAIATFPNCSKNYPCOGIVLDRVNI-KG--- 411

```

RESULT 10			
081246			
ID	081246	PRELIMINARY;	PRT; 461 AA.
AC	081246;		
DT	01-NOV-1998 (TREMBLrel. 08,	Created)	
DT	01-NOV-1998 (TREMBLrel. 08,	Last sequence update)	
DT	01-MAR-2004 (TREMBLrel. 26,	Last annotation update)	
DE	Polygalacturonase precursor.		
GN	Name=MPG3;		
OS	Cucumis melo (Muskmelon).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Cucurbitales; Cucurbitaceae; Cucumis.		
OX	NCBI_TaxID=3656;		
RN	[1]_TaxID=3656;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98289082; PubMed=9635669; DOI=10.1104/jp.117.2.363;		
RA	Patfield K.A., Rose J.K., Yaver D.S., Berks R.M., Bennett A.B.:		
RT	"Polygalacturonase precursor from Cucumis melo L. cv. 'Muskmelon'.";		

RT for polygalacturonase in ripening-associated pectin disassembly.";
 RL Plant Physiol. 117:363-373(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hadfield K.H., Rose J.K.C., Bennett A.B.;
 RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
 DR EMBL; AF062467; AAC26512.1; -
 DR PIR; T08215; T08215.
 DR GO; GO:0005618; C:cell wall; IEA.
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0004650; F:polygalacturonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR00743; Glyco_hydro_28.
 DR InterPro; IPR006626; Pbh1.
 DR InterPro; IPR011050; Pectin_lyase_like.
 DR InterPro; IPR00408; Reg_ch1_condens.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR SMART; SM00710; Pbh1; 4.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
 DR Cell wall; Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1 34 Potential.
 FT CHAIN 35 461 polygalacturonase.
 FT SEQUENCE 461 AA; 50697 MW; 358DBCB037C022B CRC64;
 Query Match 47.2%; Score 1128; DB 2; Length 461;
 Best Local Similarity 50.2%; Pred. No. 1.1e-70;
 Matches 224; Conservative 79; Mismatches 129; Indels 14; Gaps 7;
 QY 7 SILLIIIFASISICRSNVID-DNLFKQYVDNILEQFAHDFQAYLSKNIESNNMI 65
 DB 17 TFLFVVVNNEDTSTGCGSPVDNPLPSISSGDKEXYGGHYHSSSESSMLRT--RL 74
 QY 66 DKVDKNGI--KVINLISFGAKDGKTYDNIAFEQANNEACSSRTPOVYVPPKKNYLLK 122
 DB 75 EKVVSPPLASPEIFPNVDYDGMGDGED-DTEAFETWKDACSSTNAI-FLVPCDRVYHLK 132
 QY 123 QIFPSGPRSSISVKIFGSLASKSIDY-KDRRLWIAFSPVQNLVVGSGGTINGNQV 161
 DB 133 PTFPSGSPPLFLQLEGTVKASPHISDYEDKRRHWIIFQNLGLRVEHKGIIINGNRKV 192
 QY 182 WPSGCKINKSLPCRDPTALTTPNNCKLKYNNLKSNAQOIHIFESCTNVVNSMLTNA 241
 DB 193 WLNCKVKNKLPCKEAPLATVFCYCNLRKRGFRRAQOMLSFQCNVVKALNMTYA 252
 QY 242 SAKSPNTDGVHVNTOYIQTIDTIIIGTDICISIVSGSNVQATNITGFGHGISIGSLG 301
 DB 253 PGNSPMTDGIHVGTQFIVIKNCLIMTGDICISIVSGSKVRAKGIICGPHGISIGSLG 312
 QY 302 SGNSEAVSVNVTNEAKITGAENGVRKKTWOGSGGASNKFLNVENQDVKPIIIDQNY 361
 DB 313 AGSSEAVSNVNVDTAKFSGTISNGVRKKTWOGSGGAYQNIIFQNIWADNTNPITINQNY 372
 QY 362 CDREVERCIQOFSAVQVNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENILV----- 416
 DB 373 CDQKECTQADAVANVSNVYONIRGTSASAEVAVKPCSGSVCCGILLIDINLVHKGN 432
 QY 417 GEGSKPESEATCKVHFENNAEHTPHC 442
 DB 433 DDKSAQAEASCKVKKMNRGRVSPQC 458
 RESULT 11
 ID 023147 PRELIMINARY; PRT; 431 AA.
 AC 023147;
 DT 01-JAN-1998 (TEMBLrel. 05, Last Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
 DE Endo-polygalacturonase (A03957510).
 GN Name=ADPgl; Synonym=Ac3957510/T8H10_110, T8H10_110;
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21377431; PubMed=11485203; DOI=10.1023/A:1010619002833;
 RA Sander L., Child R., Ulvakov P., Albrechtsen M., Boxkhardt B.;
 RT Analysis of a dehiscence zone endo-polygalacturonase in oilseed rape
 RT (Brassica napus) and Arabidopsis thaliana: evidence for roles in cell
 RT separation in dehiscence and abscission zones, and in etylar tissues
 RT during pollen tube growth.";
 RL Plant Mol. Biol. 46:469-479(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Benes V., Rechmann S., Borkova D., Ansgore W., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X., Queciter F., Salamoubat M.;
 RL Submitted (Nov-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Jenkins E.S., Paul W., Craze M., Whitelaw C.A., Weigand A.,
 RA Roberts J.A.;
 RT "Dehiscence-related expression of an Arabidopsis thaliana gene
 RT encoding a polygalacturonase in transgenic plants of Brassica napus.";
 RL Plant Cell Environ. 22:159-167(1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Iehida J.,
 RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K. to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Shim P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
 RA Chan M.W., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,
 RA Iehida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
 RA Shinozaki K., Davis R.W., Theologis A., Becker J.R.;
 RL Submitted (Mar-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
 DR EMBL; A0002532; CA05525.1; -
 DR EMBL; AL133248; CAB66108.1; -
 DR EMBL; AF037367; AAC98923.1; -
 DR EMBL; AK117942; BAC42580.1; -
 DR EMBL; BT005376; AAO63440.1; -
 DR PIR; T46187; T46187.
 DR GO; GO:0005618; C:cell wall; IEA.
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0004650; F:polygalacturonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR00743; Glyco_hydro_28.
 DR InterPro; IPR006626; Pbh1.
 DR InterPro; IPR011050; Pectin_lyase_like.
 DR InterPro; IPR00408; Reg_ch1_condens.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR SMART; SM00710; Pbh1; 5.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
 DR Cell wall; Glycosidase; Hydrolase.
 FT SEQUENCE 431 AA; 46572 MW; 0F0F30DF45804FE7 CRC64;
 Query Match 47.2%; Score 1127; DB 2; Length 431;
 Best Local Similarity 49.6%; Pred. No. 1.1e-70;
 Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;
 QY 7 SILLIIIFASISICRSNVID-DNLFKQYVDNILEQFAHDFQAYLSYL 55
 DB 17 TFLFVVVNNEDTSTGCGSPVDNPLPSISSGDKEXYGGHYHSSSESSMLRT--RL 74

```

Db          9 AVFLCVLMLSLCKALSSNVDDGVEHDSFESDILKLNDVLT-----SLISSD 59
QY          56 SKNESNNNDKXDKNGIKYINVLSTFGAKGDKTYDNIAFEQANWACSSRTPVQFVVPK 115
Db          60 ETLLEAS-----TVSVNFGAKGDKGTDPTQAFKAWKKAACSTNGVTFLEVPK 107
QY          116 NKNYLLKQITFGSPCRSSISVKIFGSLBASKSISYKDRRLMIADSVQNLVGGG--GT 173
Db          108 GKTYLLKSTRFRPCSLNFPQILGLSLSTRKSDYKDKNMHLLIEDNNLSIDGSGTGI 167
QY          174 INNGGVWMPSSCKINKSLPCRDAPTLATFMNCKNLKVNNLKRNAAQOIHKFESCTNVV 233
Db          168 INNGKTMWNSCKIKDKSPCKTAPTLATLVLKVLNKNLKVNAQOIISTEKKNVE 227
QY          234 ASNLMLNNAKSPNTDGVHVSNTQYIOISDTIIGDDCISIVSGSQNVQATNITGPGH 293
Db          228 VSNVEITVAGDSPNTDGIHNTQINIRNSNDIGTGDICISIEGTQNIQIPLTCGPGH 287
QY          294 GISIGSLGSGSEAVSVTVNEAKIIGAENGVRIKTWGGSGQASNIKFLANVEMQVRY 353
Db          288 GISIGSLGDDNSKAYVSGINVDGAKFESDNGVRIKTYGGSGTAKNIKFNIRHENVXN 347
QY          354 PIIDONTCDRVEPCIOQFSAVQVKNVYENIKGSTATKALFPCDSTNPFCEGIIMENI 413
Db          348 PIIDODYCDK-DKCEPDQESAVQVKNVYKNISGTSATDVALTLNCSERYPCOGIVLENV 406
QY          414 NLVSGSGKPSSEATCKNVHFNNAEHVTPHCT 443
Db          407 KIKG-----GTASCKNANVKNQGVSPKCS 411

```

RESULT 12

```

ID          022935 PRELIMINARY; PRT; 426 AA.
AC          022935;
DT          01-JAN-1998 (TREMBLrel. 05, Created)
DT          01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT          01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE          Putative polygalacturonase.
CN          Name=At2g941850;
OS          Arabidopsis thaliana (Mouse-ear cress).
OC          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC          NCBI_TaxID=3702;
RN          [1]
RA          SEQUENCE FROM N.A.
RA          Rounsley S.D., Ian X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA          Spriggs T.A., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R.,
RA          Venter J.C.;
RL          Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RP          [2]
RP          SEQUENCE FROM N.A.
RP          Town C.D., Kaul S.;
RL          Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC          -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
DR          EMBL; AC002339; AAC02763.1; -.
DR          PIR; H84846; H84846.
DR          GO; GO:0005618; C:cell wall; IEA.
DR          GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR          GO; GO:0004650; F:polygalacturonase activity; IEA.
DR          GO; GO:0005975; F:carbohydrate metabolism; IEA.
DR          InterPro; IPR000743; Glyco_hydro_28.
DR          InterPro; IPR006626; PbH1.
DR          InterPro; IPR011050; Pectin_lyas_like.
DR          InterPro; IPR000408; Reg_chit_condens.
DR          Pfam; PF00295; Glyco_hydro_28; 1.
DR          SMART; SM00710; PbH1_5.
DR          PROSITE; PS00502; POLYGALACTURONASE; 1.
DR          PROSITE; PS00626; RCCL_2; UNKNOWN 1.
DR          Cell wall; Glycosidase; Hydrolase-.
DR          SEQUENCE 426 AA; 45924 MW; F756754C0390A80B CRC64;

```

```

Query Match 46.7%; Score 1115.5; DB 2; Length 426;
Best Local Similarity 49.8%; Pred. No. 7.2e-70;
Matches 222; Conservative 74; Mismatches 117; Gaps 7;
QY          6 NSILLILIIIPASSISTGCSNVVIDNLEFKQY-----DNILEQFAFDFOAYLSKMLE 60
Db          6 NLVTFELMLALMPSWCKASRISPNVYDHSYKRFKSDSLIKR--EDIGLASFPAASLR 63
QY          61 SNNIDKVDKNGIKYINVLSTFGAKGDKTYDNIAFEQANWACSSRTPVQFVVPKRYL 120
Db          64 TPT-----TVSDPFGAKGDKGTDPTQAFVNAWKKAACSSNAGVNLVPGKNTYL 112
QY          121 LKQITFGSPCRSSISVKIFGSLBASKSISYKDRRLMIADSVQNLVGGG--GTINNG 178
Db          113 LKSIQTLGPNCSILTVQIFGTLASQKRDYDIKWMIFDGVNLSLDGDDTGVVDNG 172
QY          179 QVWMPSSCKINKSLPCRDAPTLATFMNCKNLKVNNLKRNAAQOIHKFESCTNVVSNLM 238
Db          173 ETWMONSKRNKA-----KALTFYNSKSLVKNLKVRNAQOIISTEKSNVQVSNV 225
QY          239 INNAKSPNTDGVHVSNTQYIOISDTIIGDDCISIVSGSQNVQATNITGPGHSTIG 298
Db          226 VTPADSPNTDGIHNTQINIRVSESIIGTDDCISIESGQNVQINDITGPGHSTIG 285
QY          299 SLGSGNSEAVSVTVNEAKIIGAENGVRIKTWGGSGQASNIKFLANVEMQVRYPIID 358
Db          286 SLGDDNSKAFVSGTVVDGAKLSTGDNVGRIKTYGGSGTASNIIFQNIQNDVKNPIIID 345
QY          359 ONYCDRVEPCIOQFSAVQVKNVYENIKGSTATKALFPCDSTNPFCEGIIMENI 418
Db          346 QDYCDK-SKCTTESAVQVKNVYRDISTGSASNAITFTCSKNYCCGIVLDRVNIKG- 403
QY          419 SKPSEATCKNVHFNNAEHVTPHCTS 444
Db          404 ----GKATCTNANVVDKGVLPQNS 425

```

RESULT 13

```

ID          042399 PRELIMINARY; PRT; 433 AA.
AC          042399;
DT          01-NOV-1996 (TREMBLrel. 01, Created)
DT          01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT          05-JUL-2004 (TREMBLrel. 27, last annotation update)
DE          Endo-polygalacturonidase (polygalacturonase).
OS          Brassica napus (Rape).
OC          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC          NCBI_TaxID=3708;
RN          [1]
RA          SEQUENCE FROM N.A.
RA          Sander L., Bortemann J., Ulvskov P., Borkhardt B.;
RT          "Nucleotide sequence of a gene encoding a pod dehiscence zone specific
RT          endo-polygalacturonase (Accession No. X98373) from Brassica napus
RT          /F0636-0561."
RL          Plant Physiol. 111:1354-1354(1996).
RP          [2]
RP          SEQUENCE FROM N.A.
RP          MEDLINE=96382421; PubMed=8790285;
RA          Petersen M., Sander L., Child R., van Onckelen H., Ulvskov P.,
RA          Borkhardt B.;
RT          "Isolation and characterization of a pod dehiscence zone specific
RT          polygalacturonase from Brassica napus."
RL          Plant Mol. Biol. 31:517-527(1996).
CC          -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
DR          EMBL; X98373; CAA67020.1; -.
DR          EMBL; X95800; CAA67020.1; -.
DR          GO; GO:0005618; C:cell wall; IEA.
DR          GO; GO:0005978; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR          GO; GO:0004650; F:polygalacturonase activity; IEA.
DR          GO; GO:0005975; F:carbohydrate metabolism; IEA.
DR          InterPro; IPR000743; Glyco_hydro_28.

```


DR InterPro: IPR006626; Pbh1.
 DR InterPro: IPR011050; Pectin lyase like.
 DR InterPro: IPR000408; Reg_chit_condens.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR SMART: SM00710; Pbh1; 6.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 DR Cell wall; Glycosidase; Hydrolase.
 KM Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 433 AA; 46595 MW; 1EDB29ACB958260 CRC64;

Query Match 46.6%; Score 1113; DB 2; Length 433;

Best Local Similarity 49.9%; Pred. No. 1; le-69;

Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

QY 7 SILLIIIPASSISTGRSNVIDNLFKQVYDNLLEQFAHDPQAYLSYLSKNIENNNID 66
 DB 9 AVELCVLLMLACQALSNV-DDG-----YGHEDG--SFEDISLTKLNNDD 51
 QY 67 KV-----DKNGIKVINLVSPGAKGDKTYDNIAPFOAMNPAQSSRTPVQFVPEKNK 118
 DB 52 DUTLTKSDRPTTESSTVSINFGAKGDKTDTQAFKMKAKACSTNGVTTFLIRKGT 111
 QY 119 YLLKQITFGSPCRSSISVKIFGSLKASISDY-KDRRLWIAFDSYQNLVVGSG--GTIN 175
 DB 112 YLLKSIKRFQGPCKSLRSFQILGTLASATKSDYSNDKNHMLLEDVNNLSIDGSAQIVD 171
 QY 176 GNGQVWPSPSCCKINKSLPCRDAPLTALTFNNCKULKNNLKSQAQOIHKFESCTNVAS 235
 DB 172 GNGNIMWONSCKIDKSPCTKAPTALTLYMLKNLVNLRVNAQOIQISIEKCNVGVK 231
 QY 236 NLMINASAPSPNDGVHVSNTQYIOISDTIIGTGDCISIVSGSNVQATNIRCGPHGI 295
 DB 232 NKAITPGDSPTDGIHIVATKNIRISNSDIGTDCISIEDSQNVQINDLTCGPHGI 291
 QY 296 SIGSLGSGNSEAYSVNTVNEAKIIGAENGVRKIKTQSGSGQASNIKPLAVENQDVKYP 355
 DB 292 SIGSLGDSNKAIVSGIDVQATLSETDNGVRIKTYQSGSGTAKNIKFOQIRMDVKNPI 351
 QY 356 IIDQNVCDRVEPTIOQPSANQVKNVYENIKGTSATVAIKFPCSTNPFCEGIMENINL 415
 DB 352 IIDQNVCDK-DKCEQESAVQNNVYQNIKTSATDVAIMFNCISYVYPCQIGIVLENNI 410
 QY 416 VESGKPESEATCKNHFNNAEHTPHC 442
 DB 411 KG-----GRASCENNVKDKTIVSPKC 432

RESULT 14

Q43063 PRELIMINARY; PRT; 458 AA.

AC Q43063; 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Polyalacturonase (EC 3.2.1.15).
 OS Prunus persica (Peach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid 1; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=3760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee E., Speirs J., Gray J., Brady C.J.;
 RT "Homologues to the tomato endopolgalacturonase gene in the peach
 genome";
 RL Plant Cell Environ. 13:513-521(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINB=94302167; PubMed=8029352; DOI=10.1104/pp.105.1.125;
 RA Leister D.R., Speiers J., Orr G., Brady C.J.;
 RT "Peach (Prunus persica) endopolgalacturonase - cDNA isolation and mRNA
 analysis in melting and nonmelting peach cultivars";
 RL Plant Physiol. 105:225-231(1994).

RN [3]
 RP SEQUENCE FROM N.A.

RA Speirs J.;
 RL Submitted (APR-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
 DR EMBL: X77231; CAA54448.1; -.
 DR PIR: S71523; S71523.
 DR GO: GO:0016798; P:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO: GO:004650; P:polylacturonase activity; IEA.
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro: IPR00743; Glyco_hydro_28.
 DR InterPro: IPR006626; Pbh1.
 DR InterPro: IPR011050; Pectin lyase like.
 DR InterPro: IPR000408; Reg_chit_condens.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR SMART: SM00710; Pbh1; 6.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 KM Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 458 AA; 49716 MW; E7BABEC8B8B8A5 CRC64;

Query Match 46.4%; Score 1108; DB 2; Length 458;

Best Local Similarity 48.9%; Pred. No. 2; 7e-69;

Matches 226; Conservative 82; Mismatches 126; Indels 28; Gaps 8;

QY 1 MYIQRNLSILLIIIPASSISTGRSNVIDNLFKQVYDNLLEQFAHDPQAYLSYLS---- 56
 DB 1 MALQKHLVLFYVVSFCASCSYSGQEVNLSLSPVDH--EKESGYNRAHPNMNTIEG 58
 QY 57 -----KNIESNNIDKV---DKNGIKVINLVSPGAKGDKTYDNIAPFOAMNEAC 103
 DB 59 VKRMEFIKRAQOLFSEKRLERAGSKSSSVKTSVANFGAKGAD-DTFAPEKAAAC 117
 QY 104 SERTPVQVPPKKNYLLKQITFGSPCRSSISVKIFGSLKASISDYKDRRLWIAFDSV 163
 DB 118 SSGAGTALVVPQ-KTLVAPRIEISGPKSHLTMOIYIGTIEASDRSVYKDVTHLIFDNV 176
 QY 164 QNLVVGSGGTINGQGVWPPSSCKINKSLPCRD-APTALTFFNNCKULKNNLKSQAQOI 222
 DB 177 QSLVVGPGITNGNGKRMHNSCKRKPQPCQATATYTFNNCKNLVYKNLKIQAQGM 236
 QY 177 QSLVVGPGITNGNGKRMHNSCKRKPQPCQATATYTFNNCKNLVYKNLKIQAQGM 236
 QY 223 HKFESCTNVVNASNLMINASAPSPNDGVHVSNTQYIOISDTIIGTGDCISIVSGSNV 282
 DB 237 HVRFGQCKVVAEHLVTAPEDSPNTDGIHTVTKNITISSVIGTGDCISIVSGSNV 296
 QY 283 QATNITCGPHGISIGSLGSEAYSVNTVNEAKIIGAENGVRKIKTQSGSGQASNIK 342
 DB 297 QATDITCGPHGISIGSLGEDNANDHVGVFVNGAKISGTSNGVRKIKTQSGSGASNI 356
 QY 343 PLNVENQDVKYPITIDQNC-D-RVEPCIOQPSANQVKNVYENIKGTSATVAIKFPCST 401
 DB 357 FQVVENQDVTPITIDQNC-DHKNKDCITRORSVQVKNVYQNIKRTGTSATDITNCSQ 416
 QY 402 NFPCGEGIMENINLVGESGKPESEATCKNHFNNAEHTPHCT 443
 DB 417 SVPCQIGIVQNTQL-----QNAKACNNVKNPVKGVSPRCS 453

RESULT 15

Q8VXT3 PRELIMINARY; PRT; 433 AA.

AC Q8VXT3; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Endo polyalacturonase.
 OS Brassica rapa (Turnip).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid 1; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=51350;
 RN [1]
 RP SEQUENCE FROM N.A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 06:59:05 ; Search time 484 Seconds

(without alignments)
1102.849 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390
Sequence: 1 MVIGRSILLILIIIFASIS.....VTPHCTSLISEDEALLNY 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 116806243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	457	1 PCT-US03-37406-2	Sequence 2, Appl1
2	2390	100.0	457	22 US-09-791-537-50958	Sequence 50958, A
3	2390	100.0	457	26 US-10-018-604-2	Sequence 2, Appl1
4	2390	100.0	457	32 US-10-691-374-2	Sequence 2, Appl1
5	2382	99.7	457	1 PCT-US03-37406-48	Sequence 48, Appl1
6	2382	99.7	457	1 PCT-US03-37406-50	Sequence 50, Appl1
7	2382	99.7	457	32 US-10-691-374-48	Sequence 48, Appl1
8	2382	99.7	457	32 US-10-691-374-50	Sequence 50, Appl1
9	2338.5	97.8	456	1 PCT-US02-28066-1	Sequence 1, Appl1
10	2338.5	97.8	456	22 US-09-791-537-1102966	Sequence 102966, A
11	2338.5	97.8	456	30 US-10-487-753-1	Sequence 1, Appl1
12	1295.5	54.2	467	22 US-09-791-537-49108	Sequence 49108, A
13	1185	49.6	438	22 US-09-791-537-9525	Sequence 9525, Ap
14	1153	48.2	453	22 US-09-791-537-83144	Sequence 83144, A
15	1142	47.8	462	22 US-09-791-537-75907	Sequence 75907, A
16	1136	47.5	462	22 US-09-791-537-85713	Sequence 85713, A
17	1128	47.2	461	22 US-09-791-537-140414	Sequence 140414, A
18	1127	47.2	431	19 US-09-513-996A-51400	Sequence 51400, A
19	1127	47.2	431	22 US-09-791-537-36586	Sequence 36586, A
20	1127	47.2	431	33 US-10-739-930-6074	Sequence 6074, Ap
21	1127	47.2	463	19 US-09-513-996A-51399	Sequence 51399, A
22	1119.5	46.8	415	19 US-09-513-996A-51401	Sequence 51401, A
23	1115.5	46.7	426	21 US-09-708-427-16483	Sequence 16483, A
24	1115.5	46.7	426	22 US-09-791-537-7224	Sequence 7224, Ap
25	1113.5	46.6	409	21 US-09-708-427-16484	Sequence 16484, A
26	1113.5	46.6	433	22 US-09-791-537-115275	Sequence 115275, A
27	1108	46.4	458	22 US-09-791-537-17408	Sequence 17408, A
28	1103	46.2	433	20 US-09-646-679-32	Sequence 32, Appl
29	1103	46.2	433	22 US-09-791-537-101225	Sequence 101225, A
30	1103	46.2	433	33 US-10-739-930-7092	Sequence 7092, Ap
31	1103	46.2	433	33 US-10-787-968-32	Sequence 32, Appl
32	1091	45.6	460	22 US-09-791-537-52113	Sequence 52113, A
33	1061.5	44.4	335	22 US-09-791-537-70434	Sequence 70434, A
34	1058.5	44.3	404	22 US-09-791-537-17410	Sequence 17410, A
35	1011	42.3	393	30 US-10-437-963-127090	Sequence 127090, A
36	1010	42.3	452	30 US-10-437-963-127091	Sequence 127091, A
37	988.5	41.4	444	30 US-10-437-963-111044	Sequence 111044, A
38	940	39.3	340	30 US-10-425-115-218415	Sequence 218415, A
39	939.5	39.3	433	37 US-60-324-109-20499	Sequence 20499, A
40	925	38.7	452	1 PCT-US00-26884-22	Sequence 22, Appl
41	925	38.7	452	18 US-09-410-191-22	Sequence 22, Appl
42	925	38.7	452	22 US-09-791-537-54401	Sequence 54401, A
43	911	38.1	438	30 US-10-449-907-36629	Sequence 36629, A
44	898.5	37.6	459	22 US-09-791-537-123232	Sequence 123232, A
45	898.5	37.6	459	24 US-09-935-625-649	Sequence 649, App

ALIGNMENTS

RESULT 1
PCT-US03-37406-2

Sequence 2 Application PC/TUS0337406

GENERAL INFORMATION:
APPLICANT: McCallum, Claire
APPLICANT: Slade, Ann J.
APPLICANT: Colbert, Trent
APPLICANT: Knaut, Vic
TITLE OF INVENTION: Transgenic Mutations in the Polygalacturonase Gene

FILE REFERENCE: MBHB 02-276
CURRENT APPLICATION NUMBER: PCT/US03/37406
CURRENT FILING DATE: 2003-11-21
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 457
TYPE: PRT

ORGANISM: Lycopersicon esculentum
PCT-US03-37406-2

Query Match
Best Local Similarity 100.0%; Score 2390; DB 1; Length 457;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDFOAYLSYLSKNIE 60
DB 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDFOAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKYINVLSPGAKDGKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
DB 61 SNNNIDKVDKNGIKYINVLSPGAKDGKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
QY 121 LKQITFSGPCRSSISVTKIFGSLFASSTKIDYKDRRLMIAFDSVQNLVGGGGTINNGQV 180
DB 121 LKQITFSGPCRSSISVTKIFGSLFASSTKIDYKDRRLMIAFDSVQNLVGGGGTINNGQV 180
QY 181 WMPSSCKINKSLPCRDAPALTFWNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFWNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTVNEAKIIGAENGVRITKWOGSGSOASNIKFLANVEMQDVKYPPIIDON 360
DB 301 GSGNSEAVYNSVTVNEAKIIGAENGVRITKWOGSGSOASNIKFLANVEMQDVKYPPIIDON 360
QY 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
QY 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALYNY 457
DB 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALYNY 457

```

RESULT 2

US-09-791-537-50958
Sequence 50958, Application US/09791537
GENERAL INFORMATION:

```

APPLICANT: Biomolix, Inc.
APPLICANT: Dede, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 50958
LENGTH: 457
TYPE: PRT
ORGANISM: Lycopersicon esculentum
US-09-791-537-50958

```

Query Match
Best Local Similarity 100.0%; Score 2390; DB 22; Length 457;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDFOAYLSYLSKNIE 60
DB 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDFOAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKYINVLSPGAKDGKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
DB 61 SNNNIDKVDKNGIKYINVLSPGAKDGKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
QY 121 LKQITFSGPCRSSISVTKIFGSLFASSTKIDYKDRRLMIAFDSVQNLVGGGGTINNGQV 180
DB 121 LKQITFSGPCRSSISVTKIFGSLFASSTKIDYKDRRLMIAFDSVQNLVGGGGTINNGQV 180

```

```

DB 121 LKQITFSGPCRSSISVTKIFGSLFASSTKIDYKDRRLMIAFDSVQNLVGGGGTINNGQV 180
QY 181 WMPSSCKINKSLPCRDAPALTFWNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFWNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTVNEAKIIGAENGVRITKWOGSGSOASNIKFLANVEMQDVKYPPIIDON 360
DB 301 GSGNSEAVYNSVTVNEAKIIGAENGVRITKWOGSGSOASNIKFLANVEMQDVKYPPIIDON 360
QY 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
QY 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALYNY 457
DB 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALYNY 457

```

RESULT 3

US-10-018-604-2
Sequence 2, Application us/10018604
GENERAL INFORMATION:

```

APPLICANT: DANISCO A/S
TITLE OF INVENTION: Process for the Enzymatic Modification of Pectin
FILE REFERENCE: 550-354
CURRENT APPLICATION NUMBER: US/10/018,604
CURRENT FILING DATE: 2002-09-06
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 457
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: PG enzyme
US-10-018-604-2

```

Query Match
Best Local Similarity 100.0%; Score 2390; DB 26; Length 457;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDFOAYLSYLSKNIE 60
DB 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDFOAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKYINVLSPGAKDGKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
DB 61 SNNNIDKVDKNGIKYINVLSPGAKDGKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
QY 121 LKQITFSGPCRSSISVTKIFGSLFASSTKIDYKDRRLMIAFDSVQNLVGGGGTINNGQV 180
DB 121 LKQITFSGPCRSSISVTKIFGSLFASSTKIDYKDRRLMIAFDSVQNLVGGGGTINNGQV 180
QY 181 WMPSSCKINKSLPCRDAPALTFWNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFWNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTVNEAKIIGAENGVRITKWOGSGSOASNIKFLANVEMQDVKYPPIIDON 360
DB 301 GSGNSEAVYNSVTVNEAKIIGAENGVRITKWOGSGSOASNIKFLANVEMQDVKYPPIIDON 360
QY 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420

```

Db 361 YCDRVEPCIOQFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457

RESULT 4
US-10-691-374-2

Sequence 2, Application US/10691374
GENERAL INFORMATION:
APPLICANT: McCallum, Claire
APPLICANT: Slade, Ann J.
APPLICANT: Colbert, Trent
APPLICANT: Knaf, Vic
APPLICANT: Anawah Inc.
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
FILE REFERENCE: MBHB 02-276
CURRENT APPLICATION NUMBER: US/10/691,374
CURRENT FILING DATE: 2003-10-22
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 457
TYPE: PRT
ORGANISM: Lycopersicon esculentum
US-10-691-374-2

Query Match 100.0%; Score 2390; DB 32; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.2e-229;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNIIEQEFADFOAYLSYSKNIE 60
Db 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNIIEQEFADFOAYLSYSKNIE 60
QY 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAPQAMNEACSSRTPVQFVVPKKNYL 120
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAPQAMNEACSSRTPVQFVVPKKNYL 120
QY 121 LKQITSGPCRSSISYKIFGSLASAKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180
Db 121 LKQITSGPCRSSISYKIFGSLASAKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180
QY 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNLSKNAQOIHIFESCTNNVVASNLMIN 240
Db 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNLSKNAQOIHIFESCTNNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300
Db 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300
QY 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWOGSGSOASNIKFLANVEMQVKKYPIIIDQN 360
Db 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWOGSGSOASNIKFLANVEMQVKKYPIIIDQN 360
QY 361 YCDRVEPCIOQFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420
Db 361 YCDRVEPCIOQFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457

RESULT 5
PCT-US03-37406-48
Sequence 48, Application PC/TUS0337406
GENERAL INFORMATION:
APPLICANT: McCallum, Claire
APPLICANT: Slade, Ann J.
APPLICANT: Colbert, Trent

APPLICANT: Knaf, Vic
APPLICANT: Anawah Inc.
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
FILE REFERENCE: MBHB 02-276
CURRENT APPLICATION NUMBER: PCT/US03/37406
CURRENT FILING DATE: 2003-11-21
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 48
LENGTH: 457
TYPE: PRT
ORGANISM: Lycopersicon esculentum
PCT-US03-37406-48

Query Match 99.7%; Score 2382; DB 1; Length 457;
Best Local Similarity 99.8%; Pred. No. 7.8e-229;
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNIIEQEFADFOAYLSYSKNIE 60
Db 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNIIEQEFADFOAYLSYSKNIE 60
QY 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAPQAMNEACSSRTPVQFVVPKKNYL 120
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAPQAMNEACSSRTPVQFVVPKKNYL 120
QY 121 LKQITSGPCRSSISYKIFGSLASAKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180
Db 121 LKQITSGPCRSSISYKIFGSLASAKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180
QY 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNLSKNAQOIHIFESCTNNVVASNLMIN 240
Db 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNLSKNAQOIHIFESCTNNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300
Db 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300
QY 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWOGSGSOASNIKFLANVEMQVKKYPIIIDQN 360
Db 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWOGSGSOASNIKFLANVEMQVKKYPIIIDQN 360
QY 361 YCDRVEPCIOQFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420
Db 361 YCDRVEPCIOQFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457

RESULT 6
PCT-US03-37406-50
Sequence 50, Application PC/TUS0337406
GENERAL INFORMATION:
APPLICANT: McCallum, Claire
APPLICANT: Slade, Ann J.
APPLICANT: Colbert, Trent
APPLICANT: Knaf, Vic
APPLICANT: Anawah Inc.
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
FILE REFERENCE: MBHB 02-276
CURRENT APPLICATION NUMBER: PCT/US03/37406
CURRENT FILING DATE: 2003-11-21
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 50
LENGTH: 457
TYPE: PRT
ORGANISM: Lycopersicon esculentum
PCT-US03-37406-50

Query Match 99.7%; Score 2382; DB 1; Length 457;
Best Local Similarity 99.8%; Pred. No. 7,8e-229;
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MVIORNSILLLIIIFASSISTCRSNVIDDNLFKQVYDNLLEOFADHFOAYISYSKNTIE 60
DB 1 MVIORNSILLLIIIFASSISTCRSNVIDDNLFKQVYDNLLEOFADHFOAYISYSKNTIE 60
QY 61 SNNNIDKVDKNGIKVINVLISFGAKGDKTYDNIATFOAMNEACSSRTVPQFVPPKKNKYL 120
DB 61 SNNNIDKVDKNGIKVINVLISFGAKGDKTYDNIATFOAMNEACSSRTVPQFVPPKKNKYL 120
QY 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRRLMIAFDSYQNLVVGSGGTINGNOV 180
DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRRLMIAFDSYQNLVVGSGGTINGNOV 180
QY 181 MWPSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 MWPSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIISDTIIGTGDDCISIVSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIISDTIIGTGDDCISIVSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTVAEAKIIGAENGVRIKTWOGSGGQASNIKFLNVEMQDVKYPIIIDON 360
DB 301 GSGNSEAVYNSVTVAEAKIIGAENGVRIKTWOGSGGQASNIKFLNVEMQDVKYPIIIDON 360
QY 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATVAKIFPCSTNFPCEGIMENINLVGSESG 420
DB 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATVAKIFPCSTNFPCEGIMENINLVGSESG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEBALVNY 457
DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEBALVNY 457
```

RESULT 7

```
US-10-691-374-48
; Sequence 48, Application US/10691374
; GENERAL INFORMATION:
; APPLICANT: McCallum, Claire
; APPLICANT: Slade, Ann J.
; APPLICANT: Colbert, Trent
; APPLICANT: Knauf, Vic
; APPLICANT: Anawah Inc.
; TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
; FILE REFERENCE: Transgenic Mutations in the Polygalacturonase Gene
; CURRENT APPLICATION NUMBER: US/10/691,374
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-691-374-48
```

Query Match 99.7%; Score 2382; DB 32; Length 457;
Best Local Similarity 99.8%; Pred. No. 7,8e-229;
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MVIORNSILLLIIIFASSISTCRSNVIDDNLFKQVYDNLLEOFADHFOAYISYSKNTIE 60
DB 1 MVIORNSILLLIIIFASSISTCRSNVIDDNLFKQVYDNLLEOFADHFOAYISYSKNTIE 60
QY 61 SNNNIDKVDKNGIKVINVLISFGAKGDKTYDNIATFOAMNEACSSRTVPQFVPPKKNKYL 120
DB 61 SNNNIDKVDKNGIKVINVLISFGAKGDKTYDNIATFOAMNEACSSRTVPQFVPPKKNKYL 120
QY 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRRLMIAFDSYQNLVVGSGGTINGNOV 180
DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRRLMIAFDSYQNLVVGSGGTINGNOV 180
```

```
DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRRLMIAFDSYQNLVVGSGGTINGNOV 180
QY 181 MWPSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 MWPSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIISDTIIGTGDDCISIVSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIISDTIIGTGDDCISIVSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTVAEAKIIGAENGVRIKTWOGSGGQASNIKFLNVEMQDVKYPIIIDON 360
DB 301 GSGNSEAVYNSVTVAEAKIIGAENGVRIKTWOGSGGQASNIKFLNVEMQDVKYPIIIDON 360
QY 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATVAKIFPCSTNFPCEGIMENINLVGSESG 420
DB 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATVAKIFPCSTNFPCEGIMENINLVGSESG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEBALVNY 457
DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEBALVNY 457
```

RESULT 8

```
US-10-691-374-50
; Sequence 50, Application US/10691374
; GENERAL INFORMATION:
; APPLICANT: McCallum, Claire
; APPLICANT: Slade, Ann J.
; APPLICANT: Colbert, Trent
; APPLICANT: Knauf, Vic
; APPLICANT: Anawah Inc.
; TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
; FILE REFERENCE: Transgenic Mutations in the Polygalacturonase Gene
; CURRENT APPLICATION NUMBER: US/10/691,374
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-691-374-50
```

Query Match 99.7%; Score 2382; DB 32; Length 457;
Best Local Similarity 99.8%; Pred. No. 7,8e-229;
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MVIORNSILLLIIIFASSISTCRSNVIDDNLFKQVYDNLLEOFADHFOAYISYSKNTIE 60
DB 1 MVIORNSILLLIIIFASSISTCRSNVIDDNLFKQVYDNLLEOFADHFOAYISYSKNTIE 60
QY 61 SNNNIDKVDKNGIKVINVLISFGAKGDKTYDNIATFOAMNEACSSRTVPQFVPPKKNKYL 120
DB 61 SNNNIDKVDKNGIKVINVLISFGAKGDKTYDNIATFOAMNEACSSRTVPQFVPPKKNKYL 120
QY 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRRLMIAFDSYQNLVVGSGGTINGNOV 180
DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRRLMIAFDSYQNLVVGSGGTINGNOV 180
QY 181 MWPSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 MWPSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIISDTIIGTGDDCISIVSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIISDTIIGTGDDCISIVSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTVAEAKIIGAENGVRIKTWOGSGGQASNIKFLNVEMQDVKYPIIIDON 360
DB 301 GSGNSEAVYNSVTVAEAKIIGAENGVRIKTWOGSGGQASNIKFLNVEMQDVKYPIIIDON 360
```


Query Match 97.8%; Score 2338.5; DB 30; Length 456;
Best Local Similarity 98.9%; Pred. No. 1.8e-224;
Matches 452; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
QY 1 MVIQNSILLIITPASSISTCRSNVIDDNLFKQVYDNIIEQFAHDFQAYLSYLSKNIE 60
DB 1 MVIQNSILLIITPASSISTCRSNVIDDNLFKQVYDNIIEQFAHDFQAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKIVINLSFGAKGDKGTYDNTAFEQAMNEACSSRTPOVFPVKKNKYL 120
DB 61 SNNNIDKVDKNGIKIVINLSFGAKGDKGTYDNTAFEQAMNEACSSRTPOVFPVKKNKYL 120
QY 121 LKQTFSPGCRSSISVKIFGSLBASISKIDYKORRLWIAFSDYQNLVVGSGGTINNGVY 180
DB 121 LKQTFSPGCRSSISVKIFGSLBASISKIDYKORRLWIAFSDYQNLVVGSGGTINNGVY 180
QY 181 WMPSSCKINKSLPRDAPLALTFMNCNKLKVNKLKSKNAQOIHIFKESCTNVASNLMTN 240
DB 181 WMPSSCKINKSLPRDAPLALTFMNCNKLKVNKLKSKNAQOIHIFKESCTNVASNLMTN 240
QY 241 ASAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSQNVQATNITGPGHGISTGL 300
DB 241 ASAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSQNVQATNITGPGHGISTGL 300
QY 301 GSGNSEAYSNVNTVNAKIIIGAENGRIKTMQGGSGQASNIKFLVEMQDVYKPIIIDQN 360
DB 301 GSGNSEAYSNVNTVNAKIIIGAENGRIKTMQGGSGQASNIKFLVEMQDVYKPIIIDQN 360
QY 361 YCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGSSG 420
DB 361 YCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGSSG 420
QY 421 KPESEATCKVHFNMAEHVPHCTSLSEISDEALYNY 457
DB 421 KPESEATCKVHFNMAEHVPHCTSLSEISDEALYNY 456
```

RESULT 12

US-09-791-537-49108
Sequence 49108, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
CURRENT APPLICATION NUMBER: US/09/791.537
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 49108
LENGTH: 467
TYPE: PR1
ORGANISM: Actinidia chinensis
US-09-791-537-49108

Query Match 54.2%; Score 1295.5; DB 22; Length 467;
Best Local Similarity 54.6%; Pred. No. 6.7e-120;
Matches 257; Conservative 75; Mismatches 112; Indels 27; Gaps 7;

```
QY 1 MVIQNR--SILLIITPASSISTCRSNVIDDNLFKQVYDNIIEQFAHDFQAYLSYLSKN 58
DB 1 MVIQNRFPQFVITTLIPSFILGYTSVAHED---PPHDYHLE-EGYDPAKAYPSYITTI 55
QY 59 IESNNNIDKVDKNGI-----KVINLSFGAKGDKGTYDNTAFEQAMNEACSSRTPOV 101
DB 59 GNDPFGSSKSHENGIFGKRVKVDYGMDRVLDASKTYVNVDPFGAKGGRD-DTKAEKKA 114
QY 102 ACSSRTPOVFPVKKNKYLKQITSGPCRSISVYKIFGSLBASISKIDY-KDRRLWIAF 160
DB 115 ACSSRTSVAVLVFK-KNYLVRIPIPSGCKGLTQWQIYGTLEASDDSDYKDKORHMLV 173
```

```
QY 161 DSVQNLVVGSGGTINNGVQVWMPSSCKINKSLPRDAPLALTFMNCNKLKVNKLKSKNAQ 220
DB 174 DSVQNLVVGSGGTINNGVQVWMPSSCKINKSLPRDAPLALTFMNCNKLKVNKLKSKNAQ 223
QY 221 QIHIFKESCTNVASNLMTNIAKSPKPTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSQ 280
DB 224 QIHVSPDCVNVQASNLMTNIAKSPKPTDGIHTGQNIHISCVIGTGDDCISIVSGSR 293
QY 293 NVOATNITGPGHGISTGSGNSEAYSNVNTVNAKIIIGAENGRIKTMQGGSGQASNI 340
DB 294 KRVYNDITGPGHGISTGSGNSEAYSNVNTVNAKIIIGAENGRIKTMQGGSGQASNI 353
QY 341 IKFLVEMQDVYKPIIIDQNYCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFDCS 400
DB 354 IKFQVEMHNVENPIIIDQNYCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFDCS 413
QY 401 TNPCEGIMENINLVGSGKPESEATCKVHFNMAEHVPHCTSLSEISDE 451
DB 414 KRPFCQIVLEVDYDLIEBGAAMAKALCNVVELSEITGVVSPHCQEBGEBEE 464
```

RESULT 13

US-09-791-537-9525
Sequence 9525, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
CURRENT APPLICATION NUMBER: US/09/791.537
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9525
LENGTH: 438
TYPE: PR1
ORGANISM: Arabidopsis thaliana
US-09-791-537-9525

Query Match 49.6%; Score 1185; DB 22; Length 438;
Best Local Similarity 51.9%; Pred. No. 7.2e-109;
Matches 230; Conservative 76; Mismatches 121; Indels 16; Gaps 5;

```
QY 8 ILLIITPASSISTCRSNV---IDNLFQO---VYDNIIEQFAHDFQAYLSYLSKNIES 61
DB 5 ILLSVFLTLFLPSCFSSYEPFNHRDDLFWSSNVYETNRHQHGNH-----TRNSHL 56
QY 62 NNNIDKVDKNGIKIVINLSFGAKGDKGTYDNTAFEQAMNEACSSRTPOVFPVKKNKYL 121
DB 57 KNRGVAPRSPSPFPVNTFPAKAKAND-DSKAFKAMWAAACSTGIYIYAPKRRDYML 115
QY 122 KQTFSPGCRSSISVKIFGSLBASISKIDYKORRLWIAFSDYQNLVVGSGGTINNGVY 181
DB 116 KAVTFSPGCRSSISVKIFGSLBASISKIDYKORRLWIAFSDYQNLVVGSGGTINNGVY 175
QY 182 WMPSSCKINKSLPRDAPLALTFMNCNKLKVNKLKSKNAQOIHIFKESCTNVASNLMTN 241
DB 176 WPKSKINQULCLAPLAVTFECNNLRVSNIRLENAQOMHLTFQDCKNVKALNLMWTS 235
QY 242 SAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSQNVQATNITGPGHGISTGL 301
DB 236 PADSPMTDGIHVSQNTILIDQSIVRTDGCISIVSGSENVRAIGTICGPGHGISTGL 295
QY 302 GSGNSEAYSNVNTVNAKIIIGAENGRIKTMQGGSGQASNIKFLVEMQDVYKPIIIDQNY 361
DB 296 EDNSAAYSNVNTVNAKIIIGAENGRIKTMQGGSGQASNIKFLVEMQDVYKPIIIDQNY 355
QY 362 YCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGSSG 421
DB 356 CDRVACPEQKSAVQVSVLTKNIQGTSSRPVIAKFKVCSKNIPIRGISIMQVXKLVDTQQ 415
```



```

Oy      422 P-SEATCKNVHFNNAEHVTPHCT 443
         ||:|:|:|:|:|:|:|:|:|
Db      416 DVSKASCSNVKLDTRGNVSP LCT 438

```

```

RESULT 14
US-09-791-537-83144
; Sequence 83144; Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmox, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 151055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 83144
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Persea americana
; US-09-791-537-83144

```

Query March	48.2%	Score 1153	DB 22	Length 453
Best Local	55.6%	Pred. No. 1.2e-105		
Matches 225	Conservative 68	Mismatches 104	Indels 8	Gaps 5

Qy	2b	DNLFQOVYDNLTLEOFPAIDFOAYLUSYLSKINIESNNINIDKYDKNGIKYIVYLEFGKAGDG	87
Db	40	DORAYTTYGPILD--EFFSIMGPEESILS--LERNPVGPEFETPBDIDISVDDPGARGDG	96
Qy	88	KTYDNIAPEOANNEACSSRTPVOVPVKNKNYLLKOITFSGPCRSSISVKIPGSLKSSK	147
Db	97	-FDTKAFPEKAMKDACSSGSLV--IYPENKUYLLKOITFSGPCSDLRVKIRGTIEASD	153
Qy	148	ISPY--KDRRLMIAPDSVONLVYGGGCTITNGGVWWSSECKINKSLPCRDAPALLTFMN	205
Db	154	QSWVGHNKNKRIEFEDISNLTLEGGGTINGGETWMOSSCKRRKSLPCKSAPALLTFRS	213
Qy	206	CKRLKYNLSEKNAQOIHKEFESTCNVVAASNLMIASAKSEBNTDGVHVSNTQYIOLSDTI	265
Db	214	CKRLIVSDLSIKDSQCMHLSFPKODVAVSNLMWYAPHSNPTDGIH*GTGRHWNVS	273
Qy	266	IGTGDDCISIVSGSONVOATNITGCPGHGISIGSLGSGNSEAVYSNTVNEAKIENG	325
Db	274	IGTGDDCISIEBGSKNVYATNITGCPGHGISIGSLGRNNSAHVSGVLVDGNNLFTTNG	333
Qy	326	VRITKQSGSGOASNIKFLNVEMDVKYPIIIDONYCDRVEPCIQGFSAVOVKNVYENI	385
Db	334	LRIKTKQSGSGAKNIKFNQIVMHNVTNPDIIDQYCDSCXKPCPEOESAVKSVVAYMNI	393
Qy	386	KGTSATKVAIKFDCSTNPCEGIIIMENINIVGSGSKPEATCKRV	430
Db	394	RGTSAEVAIVKFPDCSSPQCGIILENNILVNGGSKETTKSCSI	438

```

RESULT 15
US-09-791-537-75907
: Sequence 75907, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791.537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 75907

```

```

; LENGTH: 462
; TYPE: PRT
; ORGANISM: Persea americana
US-09-791-537-75907

```

Query Match	47.8%	Score 1142;	DB 22;	Length 462;
Best Local Similarity	55.3%	Pred. No. 1.6e-104;		
Matches 224; Conservative	67;	Mismatches 106;	Indels 8;	Gaps 5.

Qy	28	DNLFFQVYDNLLEBFAHDFQAVLYSLYSKNIENNNIDKDKNGIKVIVNLSFGAGDG	87
Db	40	DQRAYTYTGPILD-EFSSINGFEBSILS--LEBFNPVCGEPHSPTDISIDVDEFGARGDG	96
Qy	88	KTYDNTAEQAWNEACSSRTPOVQVVPKKNYLLKQITFGSPCRSSISVKI FGSLEASSK	147
Db	97	-TDDTAFFAKAKDACCSSGSVY--IVPENKNYLLKQITFGSPCKSDLRVIRIGTIEASSD	153
Qy	148	ISDY--KDRRLMTADSYQVNLVVGCGGTINNGQVWMPSSCKINKSLPCGDAFTALTFFNN	205
Db	154	QSPWGHNNKRIIEEDISNLTLEGGGTINNGETWMDSSCKRRKSLPCKSAPTALTFFNS	213
Qy	206	CKNLKNNLKSGNAQOIHIFKESCTNVVASNLMTNAAKSPNTDGVNSVTQYIQISDRI	265
Db	214	CKNLIVSDLSIDSGQKMLSPDKCDQVYASNLMTATPEHSPBNDGHIITGTQKIHWNVS	273
Qy	266	IGTGDCDISVSGSONVQATNIITCGPHGISIGSLGSGNSEBAYVSNVTNKAIRIAGNG	325
Db	274	IGTGDCDISIEBGSIMVATNITGCPHGISISGLGRNSEAHVSGLVYDGNLFTTNG	333
Qy	326	VRIKTWGGSGGASNIKEFLANVEMQDYKPIIIDQNYCDRYEPCIQFSSAVQVKNVYENT	385
Db	334	LRIKTWGGSGGAKRIKFGQIVMHNVTNPIIIDYCDKSPCEBESAVKSVNVAYMNI	393
Qy	386	KGTSATKVAIKFDGCTNFPCEGIIMENITLVGSGKSEATCKRV	430
Db	394	RGTSASEVAVAKFDCKSPCCQGYIVGNINLVGNGCKETWSSCNI	438

```
Search completed: March 24, 2005, 07:15:02
Job time : 487 secs
```

This Page Blank (uspio)

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 07:02:13 ; Search time 1256 Seconds
(without alignments)
120.472 Million cell updates/sec

Title: US-10-691-374-2
Perfect score: 2390
Sequence: 1 NVIGRNSILLIIRIFASIS.....VTPHCTLSIEDEALVNY 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US05_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1113	46.6	433	13	US-10-151-668-2
2	1103	46.2	433	16	US-10-787-958-32
3	1011	42.3	393	16	US-10-437-963-127090
4	1010	42.3	452	16	US-10-437-963-127091
5	988.5	41.4	444	15	US-10-425-114-69361
6	870	36.4	319	15	US-10-425-114-69361
7	808	33.8	508	16	US-10-437-963-154485
8	787.5	32.9	514	10	US-09-847-208-69
9	785	32.8	534	16	US-10-437-963-134257
10	739	30.9	485	16	US-10-437-963-175508
11	734	30.7	485	16	US-10-437-963-172689
12	728	30.5	443	15	US-10-424-599-235443
13	728	30.5	456	15	US-10-425-114-44707

14	724.5	30.3	398	15	US-10-362-091-4	Sequence 4, Appl1
15	723.5	30.3	492	15	US-10-424-599-284649	Sequence 284649,
16	701.5	29.4	400	16	US-10-437-963-155040	Sequence 155040,
17	700	29.3	573	15	US-10-425-114-43413	Sequence 43413, A
18	699	29.2	496	15	US-10-425-114-66151	Sequence 66151, A
19	682	28.5	571	16	US-10-437-963-159100	Sequence 159100,
20	681	28.5	737	16	US-10-437-963-162350	Sequence 162350,
21	678	28.4	457	16	US-10-437-963-146166	Sequence 146166,
22	666	27.9	393	16	US-10-437-963-116114	Sequence 116114,
23	664.5	27.8	415	15	US-10-424-599-238512	Sequence 238512,
24	660	27.6	422	15	US-10-424-599-176320	Sequence 176320,
25	653.5	27.3	385	15	US-10-425-114-41984	Sequence 41984, A
26	653.5	27.3	411	15	US-10-425-114-38472	Sequence 38472, A
27	653.5	27.3	415	15	US-10-425-114-51531	Sequence 51531, A
28	649.5	27.2	449	16	US-10-437-963-110174	Sequence 110174,
29	643.5	26.9	602	16	US-10-437-963-127089	Sequence 127089,
30	642	26.9	320	15	US-10-424-599-204806	Sequence 204806,
31	641.5	26.8	544	16	US-10-437-963-110166	Sequence 110166,
32	635.5	26.6	539	15	US-10-424-599-176466	Sequence 176466,
33	632	26.4	386	15	US-10-425-114-50981	Sequence 50981, A
34	630	26.4	410	15	US-10-425-114-38475	Sequence 38475, A
35	630	26.4	414	15	US-10-425-114-62958	Sequence 62958, A
36	630	26.4	414	15	US-10-425-114-64531	Sequence 64531, A
37	627	26.2	407	15	US-10-425-114-65565	Sequence 65565, A
38	627	26.2	410	15	US-10-425-114-46486	Sequence 46486, A
39	627	26.2	413	15	US-10-425-114-46801	Sequence 46801, A
40	627	26.2	413	15	US-10-425-114-52982	Sequence 52982, A
41	627	26.2	413	15	US-10-425-114-56708	Sequence 56708, A
42	627	26.2	415	15	US-10-425-114-46800	Sequence 46800, A
43	627	26.2	415	15	US-10-425-114-54673	Sequence 54673, A
44	627	26.2	415	15	US-10-425-114-64510	Sequence 64510, A
45	627	26.2	416	15	US-10-425-114-65564	Sequence 65564, A

ALIGNMENTS

RESULT 1
US-10-151-668-2
Sequence 2, Application US/10151668
Publication No. US20020184660A1
GENERAL INFORMATION:
APPLICANT: UIVSKOV, Peter
APPLICANT: CHILD, Robin
APPLICANT: VAN ONCKELIN, Henri
APPLICANT: PRINSEN, Els
APPLICANT: BORKHARDT, Bernard
APPLICANT: SANDER, Lilli
APPLICANT: PETERSEN, Morten
APPLICANT: BUNDEARD, POUlsen, Gert
TITLE OF INVENTION: Seed Shattering
FILE REFERENCE: 2121-0138P
CURRENT APPLICATION NUMBER: US/10/151,668
CURRENT FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US/09/051,239
PRIOR FILING DATE: 1998-09-28
PRIOR APPLICATION NUMBER: PCT/EP96/04313
PRIOR FILING DATE: 1996-10-04
PRIOR APPLICATION NUMBER: EP 95 402241.4
PRIOR FILING DATE: 1995-10-06
PRIOR APPLICATION NUMBER: EP 95 203328.0
PRIOR FILING DATE: 1995-12-08
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 433
TYPE: PRT
ORGANISM: Brassica napus
FEATURE:
OTHER INFORMATION: Strain cv. Topaz.
US-10-151-668-2

```

Query Match      46.6%; Score 1113; DB 13; Length 433;
Best Local Similarity 49.9%; Pred. No. 5.6e-91;
Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

QY 7 SILLIIIFASSISTCRSNVIDNLFKQVYDNLIEQFAHDFQAVSYLSKINSNVID 66
DB 9 AVELCVLMLAACCOALSNV-DDG-----YGHEDG---SFESDSLITKANND 51
QY 67 KV-----DKNGIKYNLVSFGAKGDKTYNTNIAFEQAMNEACSSRTVPVFPKKN 118
DB 52 DVLTKASDRPTTRESSVSISNFGAKGDKTDPTQAFKAMKKA CSTNGVTTFLLPKGKT 111
QY 119 YLLKQTFSGPCRSSISVKIFGSLKASXISDY-KDRRLMIAFDSYQNLVGGG--GTIN 175
DB 112 YLLKSIREFRGPCSKLSFQILGTLASTKSDYSNDKNMILLIEDVNNLSIDGSAAGLYD 171
QY 176 GNGQVWPSSCKINRSLPCRDAPLTALTFFNNCKKLKYNLNKSAQAQOIHIFESCTNVAS 235
DB 172 GNGNIMWQNSCKIDSKSPCTKAPTALTLYNLKNLVKNLRVNAQOIQISIEKCNVGVK 231
QY 236 NLMINASAKSPNTDGVHVSNTQYIOISDPTIIGTGDCTISVGSQNVQATNITCGPGHGI 295
DB 232 NYKITAPGDSPTDGHIVATKNIRISNDISTGDCISIEGSGQVQINDLTGPGHGI 291
QY 296 STIGSLGSGNSEAYSVNTVNEAKITGAENGVRIKTWQSGGSAQSNIKELANVEMODVXPYI 355
DB 292 STIGSLGDDSKAYVSGIDVDGATLSETDNGVRIKTYQSGGSAQSNIKELANVEMODVXPYI 351
QY 356 IIDQNYCDRVPEPCIQOFSAVQVKNVYENIKGTSTATKYAIFPDSTNPPCGIIMENINL 415
DB 352 IIDQNYCDK-DKCEQOESAIVQVNNVYQNIQTSATDVAFNFCSVKPCQIVLENNVI 410
QY 416 VQESKPSBATCKNVHFNNAEHVTPHC 442
DB 411 KG-----GKASCENVVXDKGTVSPKC 432

RESULT 2
US-10-787-958-32
; Sequence 32, Application US/10787958
; Publication NO. US20040154053A1
; GENERAL INFORMATION:
; APPLICANT: Wyatt, Paul
; APPLICANT: Roberts, Jeremy A.
; APPLICANT: Whitelaw, Catherine
; TITLE OF INVENTION: Signal Transduction Protein Involved in Plant Dehiscence
; FILE REFERENCE: 0623_089000
; CURRENT APPLICATION NUMBER: US/10/787,958
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: GB9806113.8
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-787-958-32

Query Match      46.2%; Score 1103; DB 16; Length 433;
Best Local Similarity 49.8%; Pred. No. 4.4e-90;
Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

QY 7 SILLIIIFASSISTCRSNVID-----DVLFKQVYDNLIEQFAHDFQAVSYLSKIN 61
DB 9 AVELCVLMLAACCOALSNVDDGYGHEDGSFET--DSLITKANNDVLTLSKSDRPTTES 66
QY 62 NNNIDVKNKGKIVNLSFGAKGDKTYDNIAPQAMNEACSSRTVPQVFPKKNLYL 121
DB 67 S-----TVSISNFGAKGDKTDPTQAFKAMKKA CSTNGVTTFLLPKGKTYL 114
QY 122 KQITSGPCRSSISVKIFGSLKASXISDY-KDRRLMIAFDSYQNLVGGG--GTIN 178
DB 122 KQITSGPCRSSISVKIFGSLKASXISDY-KDRRLMIAFDSYQNLVGGG--GTIN 178

```

```

DB 115 KSIRFRGPCSKLSRSPQILGTLASTKSDYSNDKNMILLIEDVNNLSIDGSAQIYDNG 174
QY 179 QVWPSSCKINKSLPCRDAPLTALTFFNNCKKLKYNLNKSAQAQOIHIFESCTNVASNL 238
DB 175 KIMWQNSCKIDSKSPCTKAPTALTLYNLNKNLRVNAQOIQISIEKCNVGVK 234
QY 239 INASAKSPNTDGVHVSNTQYIOISDPTIIGTGDCTISVGSQNVQATNITCGPGHGISIG 298
DB 235 ITAFGDSPTDGHIVATKNIRISNDIGTDCISIEGSGQVQINDLTGPGHGISIG 294
QY 299 STIGSGNSEAYSVNTVNEAKITGAENGVRIKTWQSGGSAQSNIKELANVEMODVXPYIID 358
DB 295 STIGSDNSKAYVSGINVDGATLSETDNGVRIKTYQSGGSAQSNIKELANVEMODVXPYIID 354
QY 359 QNYCDRVPEPCIQOFSAVQVKNVYENIKGTSTATKYAIFPDSTNPPCGIIMENINLVE 418
DB 355 QNYCDK-DKCEQOESAIVQVNNVYQNIQTSATDVAFNFCSVKPCQIVLENNVI 412
QY 419 SKRPSBATCKNVHFNNAEHVTPHC 442
DB 413 ---GKASCENVVXDKGTVSPKC 432

RESULT 3
US-10-437-963-127090
; Sequence 127090, Application US/10437963
; Publication NO. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(51321)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127090
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29575C.1.pep
US-10-437-963-127090

Query Match      42.3%; Score 1011; DB 16; Length 393;
Best Local Similarity 53.3%; Pred. No. 6.9e-82;
Matches 187; Conservative 57; Mismatches 105; Indels 2; Gaps 1;

QY 72 GIKYINLVSFGAKGDKTYDNIAPQAMNEACSSRTVPVFPKKNYLLKQITSGPCR 131
DB 27 GSNVFSIQSYGAHGDGHDPTKALGPTWAAACSSAKPAVLLIPKGGYIKITTTISGPK 86
QY 132 SSISVKIFGSLKASXISDYKDR--RLMIAFDSYQNLVGGGCTTNGNGQVWPSSCKIN 189
DB 87 SSISLWQNSLVASPPRSWMSKETIRHMTLLISGVTLTYTGGITIDGNGKIWMQNSCKTN 146
QY 190 KSLPCRDAPLTALTFFNNCKKLKYNLNKSAQAQOIHIFESCTNVASNNMINASAKSPNT 249
DB 147 SKLPCTEAPLTALTFFNSCKKLKVEYLYKYNVNSQOIQISVEDCTDVMVSRSLITAPETAPNTD 206
QY 250 GNVHVSNTQYIOISDPTIIGTGDCTISVGSQNVQATNITCGPGHGISIGSLGSGNSEAYV 309
DB 207 GHITRSRDVEVTDCKIKTGDCMSIEDGTENLHVKNVCGPGHGISIGSLGDNHSEAHV 266
QY 310 SNVTVEAKITGAENGVRIKTWQSGGSAQSNIKELANVEMODVXPYIIDQNYCDRVPEPC 369
DB 267 NNVTVDNVRLLYGTANAGATIKTWQSGKSAKNIVQNNVMDVNNPFIIDQNYCDSTPCK 326

```


Query Match 36.4%; Score 870; DB 15; Length 319;
 Best Local Similarity 50.8%; Pred. No. 2.3e-63;
 Matches 162; Conservative 64; Mismatches 89; Indels 4; Gaps 3;

QY 132 SSISVKTFGSLSEASKISDY--KDRRLMIAFDSYQNLVVGSGGTINGNGQWMPSSCKIN 189
 DB 1 SSVTLVYKGLTVASPNRADWSNDNRHMIWFRSIDLKTVNGGAILDGNSEKWPBHCCKIN 60
 QY 190 KSLPCRDAPLTALTWNCKNLKVNLLKSKNAOQHKESTCTNVVNASNLMTINSAKSPND 249
 DB 61 KALPCKEAPTLASHYCVDLKVEBLKTVNSQQLHMSVESANLMLRLITAPGTSPTND 120
 QY 250 CVHNSNTQYIQSDTIIGTSDDCISIVSGSQNYQATNITCGPHGISISLGSNSEAVY 309
 DB 121 GIHTRSKDVYRVDCKITGDDCMSIENGTHNLHVSKNVCGPHGISISLGDNDNSRAEV 180
 QY 310 SNTVNAKTIIGENGRIKTMOGSGQASNIKFLVEMQDYKYPPIIDONYCDREVERCI 369
 DB 181 SGITISVOLHGTTHGARIKTKYQSGGYAKDITFQMMWYDVKNPIIIDONYCDKAKFCG 240
 QY 370 QQPSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVBESGK-PSEATCK 428
 DB 241 EDSAVQVSGVVFKNIRGTTSTKDAIKMNCSENVFCGITTQNLKMQDGKGNTRSTCQ 300
 QY 429 NVHFNNAEHVTPH-CPSLE 446
 DB 301 NAKWTFEGTVRPOPCTAIK 319

RESULT 7

US-10-437-963-154485
 ; Sequence 154485, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 154485
 ; LENGTH: 508
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_54340C.1.pep
 ; US-10-437-963-154485

Query Match 33.8%; Score 808; DB 16; Length 508;
 Best Local Similarity 36.4%; Pred. No. 1.6e-63;
 Matches 173; Conservative 91; Mismatches 167; Indels 44; Gaps 7;

QY 8 ILLIITIPASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVY-----SY 54
 DB 18 VLVALLMVAAMAKTASGGGDDGAGGDH-----HDHQFQKLWNGRGADAKEDY 69
 QY 55 LSKNIESNNIDKVDKN-----GIKVINLSPGAKGDDGTINYNTAFBQ 97
 DB 70 LMMDDDDDDDEDEEBEABQVMAAKCRPRAGRVNVNDSFGAAGDGCSDTEAFLN 129
 QY 98 AWEACSSRTPVQVYVPPKKNYTLKQITFSGPCSSSISVKIFGSLSEASKISDY--KDR 155
 DB 130 AMKAGCSLNNAV-FLVPGRRYKGAARFIPCKRMILIQITIVAPDEPSEMDPASPR 188

QY 156 LMIAFDSYQNLVVGSGGTINGNGQWMPSSCKINKSLPCRDAPLTALTWNCKNLKVNLLK 215
 DB 189 LMLFSLGIAKRIQGGGLIDSGSKWMANCKIDRSFPCGAPLTALIDSCRGVSRNLK 248
 QY 216 SKNAOQHKESTCTNVVNASNLMTINSAKSPNDGVHNSNTQYIQSDTIIGTSDDCISI 275
 DB 249 LQNAQOQHLTVRSRSDVRLASVRVDSPESPNDGIHVADSTAVTIGSCRATGDDCISI 308
 QY 276 VSGSQNYQATNITCGPHGISISLGSNSEAVVNTVNEAKTIIGENGVRIKTMOGGS 335
 DB 309 SNGSFAYRMDIDCGPHGISISLGOQGAFAVVDGVSIDGARVARAQGVRIKTMOGGA 368
 QY 336 GQASNIKFLVEMQDYKYPPIIDONYCDREVERCIQOFSAVQVKNVYENIKGTSATKVAI 395
 DB 369 GYVRNVFAVVRVGDVHPVIDQFYCDATRCRKRNTSVKSGVFNRTGTARABAI 428
 QY 396 KFCSTNFPCEGIMENINLVBESGKPSATCKN--VHFNNAEHVTPHCTSLEIS 448
 DB 429 RLACSDAVPCVGIIVLSDIDLRREDGGEVQVTCNCAMGFDDG-RVSPAADCLRTS 482

RESULT 8

US-09-847-208-69
 ; Sequence 69, Application US/09847208
 ; Publication No. US20030082190A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Saxon, Andrew
 ; APPLICANT: Zhang, Ke
 ; APPLICANT: Zhu, Daocheng
 ; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
 ; TITLE OF INVENTION: 1GE-MEDIATED ALLERGIC DISEASES
 ; FILE REFERENCE: UCE67.002A
 ; CURRENT APPLICATION NUMBER: US/09/847,208
 ; NUMBER OF SEQ ID NOS: 177
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 69
 ; LENGTH: 514
 ; TYPE: PRT
 ; ORGANISM: Cryptomeria japonica (Japanese cedar)
 ; US-09-847-208-69

Query Match 32.9%; Score 787.5; DB 10; Length 514;
 Best Local Similarity 38.5%; Pred. No. 1.1e-61;
 Matches 173; Conservative 78; Mismatches 163; Indels 35; Gaps 9;

QY 8 ILLIITIPASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVYLSKNIENNNIDK 67
 DB 12 VAMQLLITMAA-----EDQSAQIMLDSIEQ-----YL-----RSNRSILRK 47
 QY 68 VD--KNGIKVINLSPGAKGDKTYDNTAFQAMWACSSRTPVQVYVPPKKNYTLKQI 124
 DB 48 VHSRHDAINIFVEYKYGAVGDKHDCTEAFSTAWQAC-KKPSAMLLVPNGKFFVNNL 106
 QY 125 TEGSPCRSSISVKIFGSLSEASKISDYKDRRLMIAFDSYQNLVVGSGGTINGNGQWMP 184
 DB 107 FFWGPOQPHFTFKVGIILAAQONPASMKNRRIWLFALTLGTTLMGKVIDGQKQMMAG 166
 QY 185 SCK-INKSLPC--RDAPLTALTWNCKNLKVNLLKSKNAOQHKESTCTNVVNASNLMTIN 241
 DB 167 QCKWNGRITCNDRPPRIKPFSTGLIQGLKLMNSPEHVLVFGNCEGVKIGISITA 226
 QY 242 SAKSPNDGVHNSNTQYIQSDTIIGTSDDCISIVSGSQNYQATNITCGPHGISISIG 301
 DB 227 PRDSPNTGIDIFPAKNPHLQKNTIGTDDCVALTGTSSNVIEDLTCGPHGISISIG 286
 QY 302 SGNSEAYVNTVNEAKTIIGENGVRIKTMOGSGQASNIKFLVEMQDYKYPPIIIDONY 361
 DB 287 RENSRAEVSIVHVAQKFIQTQNGLRITKTMOGSGMAASHIYENVENINSENPILINQFY 346
 QY 362 CDRVPCIQPSAVQVKNVYENIKGTSATKVAIKPDCSTNFPCEGIMENINLVBESGK 421
 DB 347 CTSASACQNGRSANVQIQDYVYKNIKRGTSATMAAILQKCSDSMPCKDIKLSISLKLTSGK 406

OTHER INFORMATION: Clone ID: PAT_MRT4530_70800C.1.pep
US-10-437-963-172689

Query Match
Best Local Similarity 30.7%; Score 734; DB 16; Length 503;
Matches 153; Conservative 66; Mismatches 147; Indels 10; Gaps 6;

QY 78 VLSFGAKGDKTYNDIAFEQAMNEACSSRTPVQFVVPKNNKYLKQITFGSPGCSISV 6;
DB 114 VQDFGMAAGDVDDTDALKTAMDTCADDGAGCVLAAAGRSFLHTTTFGTGCGSVTLQ 137
QY 138 IRGSLSEASKISDY--KORRLMIAFDSQNLVVGSGGTINGQVWVWSSCKINKS---- 191
DB 174 VQGTIAPSEPAITPAANNKRMVLVFRADGVLVAGAGLIDGKGGKWMDDLCKHKGKGNH 233
QY 197 LPCRAPALTFWNCNKLKVNLLSKNAQOIHIFESGTVNVASNLMTNASASPTDGV 251
DB 234 GPC-DSPVAMRFALSNNTVTRGLKQNSPEHFHFDGNGVRVDGLSISPSALSPMTDGI 292
QY 252 HVSNTQYQISDTIIGTDCISIVSGSQNVQATNITGPGHGISIGSLGSGNSEAYVGN 311
DB 293 HVENISDVLINTVVSNGDCVSIAGATLNVHLENTVCGPHGISIGSLGKAGTACVAN 352
QY 312 VTVNEAKIENGVRIRKTMQSGSGQASNIKFLNVEMQDVKPIIIDQNYCDRVEPTQO 371
DB 353 VTVNNAVIRHSDNGVRIRKTMQSGSGSAVAFENVMDAVRNPFIIDQYCC-LSKSCENE 411
QY 372 FSAVOVKNVYENIKGTSATK-VAIKFDCSTNPPCGIIMENINLVGSGKP-SEATCKN 429
DB 412 TTVNVVGVSGIRGTIVRGPPIHFQGSDAVPCTNITLSDVELLPASGDTVEBPCKN 471
QY 430 VHRNNAEHVTHCTSL 445
DB 472 VYGNMAATPTVPVSVCL 487

RESULT 12

US-10-424-599-235443
Sequence 235443, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 235443
LENGTH: 443
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_54623C.1.pep
US-10-424-599-235443

Query Match
Best Local Similarity 30.5%; Score 728; DB 15; Length 443;
Matches 158; Conservative 51; Mismatches 139; Indels 20; Gaps 10;

QY 77 NVLSFAKDGKTYNDIAFEQAMNEACSSRTPVQFVVPKNNKYLKQITFGSP-GRSIS 135
DB 47 NVLDYGAAGDGHADDTKAFEDMAAACVVEGST-MVVPSSGVFLVKFISFGPCEPNIV 105
QY 136 VVIFGSLSEASKISDYKDRRL-WIAFDSVQNLVVGSGGTINGQVWV-----PSSC 186
DB 106 FQLDGKIAPTSSEMGSGTQMLEFSKLTNTITRGKGVIDGQSVWNNDSPTYNPTVE 165
QY 167 KI--NKSIPCRDAPALTFFWNCNKLKVNLLSKNAQOIHIFESGTVNVASNLMTNASK 244

DB 166 MLESNGRLP-STKPTALRFYSGDGVTVGTRIQNSQOQTHLKFPSCNTNVQVSGISVSSPD 224
QY 245 SPNTDGVHVSNTQYQISDTIIGTDCISIVSGSQNVQATNITGPGHGISIGSLGSGN 304
DB 225 SPNTDGHILQNSQNVVYSSITLACGDDCVSIQTGCSDDIYHANNCCPGHGISIGSLGREN 284
QY 305 SEAYSVNTVNEAKIENGVRIRKTMQSGSGQASNIKFLNVEMQDVKPIIIDQNYCDR 364
DB 295 TKACVRNVTVARDVTIQTTLTGVRIRKTMQSGSGSVQIMFSNVQVSGVQTPISIDYCDG 344
QY 365 VEPICQFSAVOYKRVNVEYENIKGTSATPKVAIKFDCSTNPPCGIIMENINLVGSGKPS 424
DB 345 GR-CRNSSSAVAVSGIHVYVVKGT-YTKEPITFACSDNLPFGSITLDTIOL--ESAQ--E 398
QY 425 ATCKNVHF 432
DB 399 TKNSNVVF 406

RESULT 13

US-10-425-114-44707
Sequence 44707, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 44707
LENGTH: 456
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700847226_FLI.pep
US-10-425-114-44707

Query Match
Best Local Similarity 42.9%; Score 728; DB 15; Length 456;
Matches 158; Conservative 51; Mismatches 139; Indels 20; Gaps 10;

QY 77 NVLSFAKDGKTYNDIAFEQAMNEACSSRTPVQFVVPKNNKYLKQITFGSP-GRSIS 135
DB 60 NVLDYGAAGDGHADDTKAFEDMAAACVVEGST-MVVPSSGVFLVKFISFGPCEPNIV 118
QY 136 VVIFGSLSEASKISDYKDRRL-WIAFDSVQNLVVGSGGTINGQVWV-----PSSC 186
DB 119 FQLDGKIAPTSSEMGSGTQMLEFSKLTNTITRGKGVIDGQSVWNNDSPTYNPTVE 178
QY 167 KI--NKSIPCRDAPALTFFWNCNKLKVNLLSKNAQOIHIFESGTVNVASNLMTNASK 244
DB 179 MLESNGRLP-STKPTALRFYSGDGVTVGTRIQNSQOQTHLKFPSCNTNVQVSGISVSSPD 237
QY 245 SPNTDGVHVSNTQYQISDTIIGTDCISIVSGSQNVQATNITGPGHGISIGSLGSGN 304
DB 228 SPNTDGHILQNSQNVVYSSITLACGDDCVSIQTGCSDDIYHANNCCPGHGISIGSLGREN 297
QY 305 SEAYSVNTVNEAKIENGVRIRKTMQSGSGQASNIKFLNVEMQDVKPIIIDQNYCDR 364
DB 298 TKACVRNVTVARDVTIQTTLTGVRIRKTMQSGSGSVQIMFSNVQVSGVQTPISIDYCDG 357
QY 365 VEPICQFSAVOYKRVNVEYENIKGTSATPKVAIKFDCSTNPPCGIIMENINLVGSGKPS 424
DB 358 GR-CRNSSSAVAVSGIHVYVVKGT-YTKEPITFACSDNLPFGSITLDTIOL--ESAQ--E 411
QY 425 ATCKNVHF 432

Db 412 TKNSVVF 419

RESULT 14
US-10-362-091-4

/ Sequence 4, Application US/10362091
/ Publication No. US20040049809A1

/ GENERAL INFORMATION:

/ APPLICANT: Instituto de Ciencia Aplicada e Tecnologia (ICAT)

/ TITLE OF INVENTION: Pear genes coding for b-galactosidase, Pectin Methyltransferase,
/ TITLE OF INVENTION: Polysaccharonase, Expansin and their use.

/ FILE REFERENCE: none

/ CURRENT APPLICATION NUMBER: US/10/362,091

/ PRIOR FILING DATE: 2003-02-12

/ PRIOR APPLICATION NUMBER: PT 102511 C

/ NUMBER OF SEQ ID NOS: 31

/ SOFTWARE: PatentIn version 3.1

/ SEQ ID NO 4

/ LENGTH: 398

/ TYPE: PRT

/ ORGANISM: Pyrus communis

/ US-10-362-091-4

Query Match

Best Local Similarity 41.4%; Pred. No. 3.5e-56; Length 398;
Matches 158; Conservative 64; Mismatches 145; Indels 15; Gaps 10;

64 NIDKVDKNGIKVNLVSPGAKGDKTYDNIAPFOAWNEACSSRTPVQFVVPKKNKYLKQ 123

23 SINVDAAV-TFSVSLGAKADGSTDTAKAFLSAMNACASVNPAYIVPAGR-FLGN 80

124 ITTSGGRC-SSISVKIFGSLPSSKISDYK---DRRLMIFDSYQNLVVGCGGTINGNQ 179

81 AVFSGGCKNNAIFRIAGTLVAP---SDYEVIGNAGWMLFQHVNGVTI-SGGVLDGQGT 136

180 VVWPSGCKINKSLPCRDAPALTFFWNCNKLKNNLKSNAQOIHIFESCTNVVASNIM 239

137 GLM-DCK-SGSGSCSGATTTLSFSNNVNVVSGLSLNSQMPHIIVNCGQKMGCVK 193

240 NASAKSPNTDGVVNTQYIOISDTIIGTDDCISIVSGSONVQATNITGPGHGISIGS 299

194 NAGNBPNTDGIHVWSSGVTILDSKISTGDDCVSGPTTNLIEWVACGPGHGISIGS 253

300 LGGGSEAYVSNVTVEAKITGAENGVRITKW-QGSGQASNIKFLNVMDQYKPIIID 358

254 LGHQOQEGAVQNTVKTVFTGTENGVRIKSGRPSTGFARSILFQHI VMTNVQNPIVID 313

359 ONYCDVPERIQPSANOVKNVYENIKTSATKVAIKFDCSTNPECEGIIIMENINLVGE 418

314 QNTCPNDKGCPCGQASGVKSDVTYQDIGHTSATEVAVKFDSCSNVPCNJRLODVKLT-Y 372

419 SGKPSKATCKNVHFNNAHYTP 440

373 NMQAASCTIHAGTTAGTVQP 394

Db

RESULT 15

US-10-424-599-284649

/ Sequence 284649, Application US/10424599

/ Publication No. US20040031072A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa Thomas J

/ APPLICANT: Kovalic David K

/ APPLICANT: Zhou Yihua

/ APPLICANT: Cao Yongwei

/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(53223)B

/ CURRENT APPLICATION NUMBER: US/10/424,599

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 285684

/ SEQ ID NO 284649

/ LENGTH: 492

/ TYPE: PRT

/ ORGANISM: Glycine max

/ FEATURE:

/ OTHER INFORMATION: Clone ID: PAT_MRT3847_99063C.1.pep

/ US-10-424-599-284649

Query Match

Best Local Similarity 39.6%; Pred. No. 5.9e-56; Length 492;
Matches 155; Conservative 64; Mismatches 157; Indels 15; Gaps 7;

75 VINLSPGAKGDKTYDNIAPFOAWNEACSSRTPVQFVVPKKNKYLKQITFGSPCRSS1 134

96 VFDVRSFGAVGDCADDTAFRAAMKAAACVDSGI-VLAPENYSFKITSTIFSGCPKPL 154

135 SVKIFGSLPSSKISDYK---RRLLMIFDSYQNLVVGCGGTINGNQVWVWPSCKINK- 130

155 VFQVDTLMAFDGPNVSWPEADSNQWLVFRLDQMTLNGTITGEGNDKWDLPCKPHRG 214

191 -----SLPCRDAPALTFFWNCNKLKNNLKSNAQOIHIFESCTNVVASNIMINSAK 244

215 PNGKTLSGPC-GSPAMIRPFMSNKLKVKGLKIQNSFOFHMIFNGCGVLIDKLSISPRK 273

245 SPNTDGVHVSNTQYIOISDTIIGTDDCISIVSGSONVQATNITGPGHGISIGSLGSGN 304

274 SPNTDGIHNSKYVGIYNSMISNGDDCISIGPSSNVDAIGLTCPSHGISIGSLGVHN 333

305 SEAYVSNVTVEAKITGAENGVRITKWQSGGQASNIKFLNVMDQYKPIIIDQNYCDR 364

334 SQACVSNLTVRDSIIRSDNGLRITKWQSGGQSVSIRFENIOMENVGNCIIIDQYVCLS 393

365 VEPICQFSAVQKNVYENIKTSATKVA-IFDCSTNPECEGIIIMENINLVG-ESGKP 422

394 KE-CLNQTSKAVHNDVYSNIKGTIVRTPIHFACSDVACTNITLSEVELLPFGALL 452

423 SEATCKNVHFNNAHYTPHCTSLSEDEAL 453

453 DDPFCMNAVGTQRTLTIPINCAREGDETV 483

Db

Search completed: March 24, 2005, 07:36:03

Job time : 1258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: March 24, 2005, 06:59:05 ; Search time 44 Seconds
(without alignments) 775.332 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MVIQRNLSILLIIFAFSSIS.....VTPHCTSLSEDEALVNY 457

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/iaa/5A COMB .pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B COMB .pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A COMB .pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B COMB .pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB .pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfile1 .pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	457	6	5447867-1 Patent No. 5447867
2	2390	100.0	457	6	5447867-1 Patent No. 5447867
3	1113	46.6	433	4	US-09-051-239A-2 Sequence 2, App1
4	1113	46.6	433	4	US-10-151-668-2 Sequence 2, App1
5	1103	46.2	433	3	US-08-941-532-6 Sequence 6, App1
6	787	32.9	514	3	US-08-467-023-134 Sequence 134, App
7	365	15.3	70	6	5453566-2 Patent No. 5453566
8	319	13.3	127	3	US-08-467-023-189 Sequence 189, App
9	299	12.5	451	4	US-09-107-532A-6652 Sequence 6652, App
10	298.5	12.5	452	1	US-08-250-978A-5 Sequence 5, App1
11	298.5	12.5	452	2	US-08-780-869-5 Sequence 5, App1
12	298.5	12.5	452	2	US-09-787-583-2 Sequence 2, App1
13	293.5	12.3	360	3	US-09-542-767A-1 Sequence 1, App1
14	257.5	10.8	440	1	US-08-061-062A-6 Sequence 6, App1
15	241	10.1	440	1	US-08-061-062A-8 Sequence 8, App1
16	241	10.1	440	1	US-08-536-150-6 Sequence 6, App1
17	241	10.1	440	3	US-08-536-150-8 Sequence 8, App1
18	226.5	9.5	127	3	US-08-467-023-188 Sequence 188, App
19	212.5	8.9	442	4	US-09-107-532A-6254 Sequence 6254, App
20	205.5	8.6	415	4	US-09-198-956-6 Sequence 6, App1
21	205.5	8.6	415	4	US-09-198-956-6 Sequence 6, App1
22	205.5	8.6	415	4	US-09-198-956-6 Sequence 6, App1
23	188.5	7.9	128	3	US-08-467-023-187 Sequence 187, App
24	187	7.8	49	3	US-08-941-532-8 Sequence 8, App1
25	147	6.2	788	4	US-09-733-643B-14 Sequence 14, App1
26	129.5	5.4	1005	4	US-09-206-942-41 Sequence 41, App1
27	129.5	5.4	1011	4	US-09-206-942-39 Sequence 39, App1

28	127	5.3	670	4	US-09-107-433-4976 Sequence 4976, App
29	127	5.3	708	4	US-09-583-110-3019 Sequence 3019, App
30	126	5.3	785	4	US-09-733-643B-2 Sequence 2, App1
31	124.5	5.2	901	3	US-09-134-001C-5351 Sequence 5351, App
32	123.5	5.2	592	4	US-09-302-626B-6 Sequence 6, App1
33	121.5	5.1	54	4	US-09-733-643B-21 Sequence 21, App1
34	120.5	5.0	1004	4	US-09-206-942-55 Sequence 55, App1
35	120.5	5.0	1010	4	US-09-206-942-55 Sequence 55, App1
36	120	5.0	1535	3	US-08-755-587-185 Sequence 185, App
37	119.5	5.0	592	4	US-09-302-626B-91 Sequence 91, App1
38	119.5	5.0	592	4	US-09-302-626B-98 Sequence 98, App1
39	119.5	5.0	592	4	US-09-302-626B-106 Sequence 106, App
40	119.5	5.0	592	4	US-09-302-626B-107 Sequence 107, App
41	118.5	5.0	1073	4	US-09-206-942-49 Sequence 49, App1
42	118.5	5.0	1079	4	US-09-206-942-47 Sequence 47, App1
43	117	4.9	378	3	US-09-134-001C-4692 Sequence 4692, App
44	117	4.9	1574	4	US-09-302-626B-179 Sequence 179, App
45	117	4.9	1978	4	US-09-302-626B-60 Sequence 60, App1

ALIGNMENTS

RESULT 1	5447867-1	100.0%;	Score 2390;	DB 6;	Length 457;
Patent No. 5447867		Best Local Similarity	100.0%;	Pred. No. 8.6e-208;	Indels 0;
APPLICANT: BRIDGES, IAN;SCHUCH, WOLFGANG;GRIERSON, DONALD		Matches 457;	Conservative 0;	Mismatches 0;	Gaps 0;
TITLE OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN					
ESTERASE GENE SEGMENTS					
NUMBER OF SEQUENCES: 4					
CURRENT APPLICATION DATA:					
APPLICATION NUMBER: US/08/24, 866					
FILING DATE: 26-FEB-1993					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER: 720, 629					
FILING DATE: 25-JUN-1991					
APPLICATION NUMBER: 419, 779					
FILING DATE: 29-SEP-1989					
APPLICATION NUMBER: 119, 614					
FILING DATE: 12-NOV-1987					
SEQ ID NO:1:					
LENGTH: 457					
5447867-1					
Query Match					
Best Local Similarity					
Matches 457;					
Conservative 0;					
Mismatches 0;					
Gaps 0;					
1 MVIQRNLSILLIIFAFSSISCTCRSNVYDDNLFQVYVDNILEQFPAHDFQAYLSYLSKNIE 60					
1 MVIQRNLSILLIIFAFSSISCTCRSNVYDDNLFQVYVDNILEQFPAHDFQAYLSYLSKNIE 60					
61 SNNNIDKVDKNGIKIVNVSFGAKGDKTYDNIAFEQAMNEACSSRTPOFVVPKKNYL 120					
61 SNNNIDKVDKNGIKIVNVSFGAKGDKTYDNIAFEQAMNEACSSRTPOFVVPKKNYL 120					
61 SNNNIDKVDKNGIKIVNVSFGAKGDKTYDNIAFEQAMNEACSSRTPOFVVPKKNYL 120					
121 LKQITSGPGRSSISVKIRGSLFASASKISPYKORRLMIARDSONVLVVGSGGTINGNGOV 180					
121 LKQITSGPGRSSISVKIRGSLFASASKISPYKORRLMIARDSONVLVVGSGGTINGNGOV 180					
121 LKQITSGPGRSSISVKIRGSLFASASKISPYKORRLMIARDSONVLVVGSGGTINGNGOV 180					
181 WPFSSCKINKSLPCRDAPFALTFFWNCNKLKNNLKSNQAQIHKFESCTNVVASNLMIN 240					
181 WPFSSCKINKSLPCRDAPFALTFFWNCNKLKNNLKSNQAQIHKFESCTNVVASNLMIN 240					
181 WPFSSCKINKSLPCRDAPFALTFFWNCNKLKNNLKSNQAQIHKFESCTNVVASNLMIN 240					
241 ASAKSPNTDGVHVSNTQYIQISDTTIGTGDCISIVSGSONVQATNITCGPHGISISGL 300					
241 ASAKSPNTDGVHVSNTQYIQISDTTIGTGDCISIVSGSONVQATNITCGPHGISISGL 300					
241 ASAKSPNTDGVHVSNTQYIQISDTTIGTGDCISIVSGSONVQATNITCGPHGISISGL 300					
301 GSGNSAYVSNVNVNNAKIIIGANGVRKIKTWOGSGGASNIKFLNVEMODVKKPIIIDON 360					
301 GSGNSAYVSNVNVNNAKIIIGANGVRKIKTWOGSGGASNIKFLNVEMODVKKPIIIDON 360					
301 GSGNSAYVSNVNVNNAKIIIGANGVRKIKTWOGSGGASNIKFLNVEMODVKKPIIIDON 360					
361 YCDRVEPCIQQFSAVQVKNVVENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGESC 420					

```

Db      361 YCDRVBPCIQOFSAVQVKNVYENIKGSAKVAIKFCSTNPGGIMENINLVGSG 420
Qy      421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLVNY 457
Db      421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLVNY 457

```

RESULT 2

```

5447867-1
; Patent No. 5447867
; APPLICANT: BRIDGES, TAN; SCHUCH, WOLFGANG; GRIERSON, DONALD
; TITLE OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN
; ESTERASE GENE SEGMENTS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/24,866
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 720,629
; FILING DATE: 25-JUN-1991
; APPLICATION NUMBER: 419,779
; FILING DATE: 29-SEP-1989
; APPLICATION NUMBER: 119,614
; FILING DATE: 12-NOV-1987
; SEQ ID NO:1
; LENGTH: 457
5447867-1

```

```

Query Match      100.0%; Score 2390; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 8,6e-208;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MVIORNSILLIIIFASSISTCRSNVIDNLFPQVYDNLLEQEFADPOAVLYSKNIE 60
Db      1 MVIORNSILLIIIFASSISTCRSNVIDNLFPQVYDNLLEQEFADPOAVLYSKNIE 60
Qy      61 SNNNIDKVDKNGIKVYNLUSFGAKGDKTYDNI AFEQAMNEACSRTPVQFVVKKNYL 120
Db      61 SNNNIDKVDKNGIKVYNLUSFGAKGDKTYDNI AFEQAMNEACSRTPVQFVVKKNYL 120
Qy      121 LKQITTSFGPCRSSISVYKIFGSL EASSKISDYKDRRLMTAFPSYONLVVGGGTTNGGOY 180
Db      121 LKQITTSFGPCRSSISVYKIFGSL EASSKISDYKDRRLMTAFPSYONLVVGGGTTNGGOY 180
Qy      181 WMPSSCKIKKSLPCGDAPTALT FPNCKNLKVNNLKSNNAQOIHKFESCTNVVASNLMIN 240
Db      181 WMPSSCKIKKSLPCGDAPTALT FPNCKNLKVNNLKSNNAQOIHKFESCTNVVASNLMIN 240
Qy      241 ASAKSPNTDGVHVSNTQYIQISDPTIIGTGDCTSI VSGSONVQATNITGPGHGISIGSL 300
Db      241 ASAKSPNTDGVHVSNTQYIQISDPTIIGTGDCTSI VSGSONVQATNITGPGHGISIGSL 300
Qy      301 GSGNSEAVNVTVNEAKITIGANGVRIKTWGGSGGQASNIKFLNVMQDVYKPIIIDON 360
Db      301 GSGNSEAVNVTVNEAKITIGANGVRIKTWGGSGGQASNIKFLNVMQDVYKPIIIDON 360
Qy      361 YCDRVBPCIQOFSAVQVKNVYENIKGSAKVAIKFCSTNPGGIMENINLVGSG 420
Db      361 YCDRVBPCIQOFSAVQVKNVYENIKGSAKVAIKFCSTNPGGIMENINLVGSG 420
Qy      421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLVNY 457
Db      421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLVNY 457

```

RESULT 3

```

US-09-051-239A-2
; Sequence 2, Application US/09051239A
; Patent No. 6420628
; GENERAL INFORMATION:
; APPLICANT: ULVSKOV, Peter
; APPLICANT: CHILD, Robin

```

```

; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els
; APPLICANT: BORKHARDT, Bernard
; APPLICANT: SANDER, Lilli
; APPLICANT: PETERSEN, Morten
; APPLICANT: BUNDGARD POULSEN, Gert
; APPLICANT: BOTTERMAN, Johan
; TITLE OF INVENTION: Seed Shattering
; FILE REFERENCE: 2121-0138P
; CURRENT FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: US/09/051,239A
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: PCT/EP96/04313
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: EP 95 402241.4
; PRIOR FILING DATE: 1995-12-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Strain cv. Topaz.
US-09-051-239A-2

```

```

Query Match      46.6%; Score 1113; DB 4; Length 433;
Best Local Similarity 49.9%; Pred. No. 3.1e-92;
Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

```

```

Qy      7 SILLIIIFASSISTCRSNVIDNLFPQVYDNLLEQEFADPOAVLYSKNIESNNID 66
Db      9 AVFLCVLIMLALCCQALSSNV-DDG-----YGHEDG---SFEPSDLIRLNND 51
Qy      67 KY-----DKNGIKVYNLUSFGAKGDKTYDNI AFEQAMNEACSRTPVQFVVKKNYL 118
Db      52 DVLTKSDPRTTSSIVSVNSFGAGDKTDPTQAFKAWKKA CSTNGVTTFLPKCT 111
Qy      119 YLLKQITTSFGPCRSSISVYKIFGSL EASSKISDYKDRRLMTAFPSYONLVVGGG--GTIN 175
Db      112 YLLKSIIRFRGPKCKLSFQILGTL SASTGRSDYSNDKHMLILBDVNNLSIDGSGAGIVD 171
Qy      176 GNGQVWMPSSCKIKKSLPCGDAPTALT FPNCKNLKVNNLKSNNAQOIHKFESCTNVVAS 235
Db      172 GNGNIMWQNSCTIDSKPCTTAPALTLYNKNLNVKLRVRNAQOIQISEKNNVGVK 231
Qy      236 NMTNNAKSPNTDGVHVSNTQYIQISDPTIIGTGDCTSI VSGSONVQATNITGPGHG1 295
Db      232 NVKITA PGDSPNTDGIHIVATKNI RISSDICTGDDCTSI EDGSONVQINDLTGCGHG1 291
Qy      296 SIGSLGSGNSEAVNVTVNEAKITIGANGVRIKTWGGSGGQASNIKFLNVMQDVYKPI 355
Db      292 SIGSLGDNNSKAYVSGIVDQATLSETDNGVRIKTWGGSGGTAKNIKFQNIIMDVKNPI 351
Qy      356 IIDQNYICRVBPCIQOFSAVQVKNVYENIKGSAKVAIKFCSTNPGGIMENINLV 415
Db      352 IIDQNYICRVBPCIQOFSAVQVKNVYENIKGSAKVAIKFCSTNPGGIMENINLV 410
Qy      416 VGESGKPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLVNY 457
Db      411 KG-----GKASCENNVVYKDGIVSPKC 432

```

RESULT 4

```

US-10-151-668-2
; Sequence 2, Application US/10151668
; Patent No. 6797861
; GENERAL INFORMATION:
; APPLICANT: ULVSKOV, Peter
; APPLICANT: CHILD, Robin
; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els

```

APPLICANT: BORKHARDT, Bernard
APPLICANT: SANDER, Lilli
APPLICANT: PETERSEN, Morten
APPLICANT: BUNDGARD, Poulsen, Gert
APPLICANT: BOTTERMAN, Johan
TITLE OF INVENTION: Seed Shattering
FILE REFERENCE: 2121-0138P
CURRENT APPLICATION NUMBER: US/10/151,668
CURRENT FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US/09/051,239
PRIOR FILING DATE: 1998-09-28
PRIOR APPLICATION NUMBER: PCT/EP96/04313
PRIOR FILING DATE: 1996-10-04
PRIOR APPLICATION NUMBER: EP 95 402241.4
PRIOR FILING DATE: 1995-10-06
PRIOR APPLICATION NUMBER: EP 95 203328.0
PRIOR FILING DATE: 1995-12-08
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 433
TYPE: PRT
ORGANISM: Brassica napus
FEATURE:
OTHER INFORMATION: Strain cv. Topaz.
US-10-151-668-2

Query Match 46.6%; Score 1113; DB 4; Length 433;

Best Local Similarity 49.9%; Pred. No. 3.1e-92; Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

7 STILLIIIFASSISTCRSNVYIDNLFKQVYDNILEQFAHDFQAYLSYLSKNIESNNID 66
9 AYLCLVLMACCOALSSNV-DDG-----YGHEDG---SFEHDSILKLNND 51
67 KV-----DKNGIKVINLVSPGAKDGKTYDNIAFEQAMNACSSRTVPQFVPRKNX 118
52 DVLTKSDRPTRESSIVSNFGAKDGKTDQAFKKAMKACSTNGVTTFILPKGKT 111
119 YLLKQTFSPGCRSSISVXIFGSLBASKSIDY-KDRRLMIADSVQNLVVGSG--GTIN 175
112 YLLKSIRFRGPKLSRFLGLTSLASSTKRSYSDNKNHMLLEDVNNLSIDGSGAGIYD 171
176 GNGQWMPSCCKINKSLPCRDAPLTALTFNNCKNLKVNLSKNAQOIHKFESCTNVVAS 235
172 GNNINWQNSCKIDSKPCTKAPLTALTYLNKNLVNNAQOIQISIEKCNVGVK 231
236 NLMINASAKSPNTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSONVQATNITCGPHGI 295
232 NVKITPAGSPNTDGIHIVATKNIRISNSDIGTDDCISIEDGSONVQINDLTCGPHGI 291
296 SIGSLGSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGGQASNTKFLNVEQDYKPI 355
292 SIGSLDDNSKAVYSGIDVAGATLSETDNGVRITKYQSGGTAKNIKIFQIRMDNVKNPI 351
356 IIDQVCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIMENINL 415
352 IIDQVCDK-DKCEQGESAVQVNNVYRNIGTSATDAIMFNCISKYKPCQGLVLENNI 410
416 VGESEATCKNVHFNNAEHTPHC 442
411 KG-----GRASCKNVNVKDKGTIVSPKC 432

RESULT 5
US-08-941-532-6
Sequence 6, Application US/08941532
Patent No. 6096946
GENERAL INFORMATION:
APPLICANT: ROBERTS, Jeremy Alan
APPLICANT: COUPE, Simon Allan
APPLICANT: JENKINS, Elizabeth Sarah
TITLE OF INVENTION: CONTROL OF POD DEHISCENCE

NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941,532
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00757
FILING DATE: 29-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9506684.1
FILING DATE: 31-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0623.0580001/RWE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-941-532-6

Query Match 46.2%; Score 1103; DB 3; Length 433;

Best Local Similarity 49.8%; Pred. No. 2.5e-91; Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

7 STILLIIIFASSISTCRSNVID-----DNLFKQVYDNILEQFAHDFQAYLSYLSKNIES 61
9 AYLCLVLMACCOALSSNVDDYGHEDGSEF--DSLKLNNDDVLTLSKSDRPTES 66
62 NNNIDRVKNGIKVINLVSPGAKDGKTYDNIAFEQAMNACSSRTVPQFVPRKNXYLL 121
67 S-----TVSVSNFGAKDGKTDQAFKKAMKACSTNGVTTFILPKGTYLL 114
122 KOTTFSPGCRSSISVXIFGSLBASKSIDY-KDRRLMIADSVQNLVVGSG--GTINGN 178
115 KSTIRFRGPKLSRFLGLTSLASSTKRSYSDNKNHMLLEDVNNLSIDGSGAGIYDNG 174
179 QVWMPSCCKINKSLPCRDAPLTALTFNNCKNLKVNLSKNAQOIHKFESCTNVVASNLM 238
175 KIMWQNSCKIDSKPCTKAPLTALTYLNKNLVNNAQOIQISIEKCNVGVK 234
239 INASAKSPNTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSONVQATNITCGPHGISIG 238
235 ITAPGDSPTDGIHIVATKNIRISNSDIGTDDCISIEDGSONVQINDLTCGPHGISIG 294
299 SIGSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGGQASNTKFLNVEQDYKPIIID 358
295 SIGDDNSKAVYSGIDVAGATLSETDNGVRITKYQSGGTAKNIKIFQIRMDNVKNPIIID 354
359 QNYCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIMENINLVGE 418
355 QNYCDK-DKCEQGESAVQVNNVYRNIGTSATDAIMFNCISKYKPCQGLVLENNIKG- 412
419 SGKSEATCKNVHFNNAEHTPHC 442
413 ---GRASCKNVNVKDKGTIVSPKC 432

RESULT 6

US-08-467-023-134
Sequence 134, Application US/08467023
Patent No. 6090386
GENERAL INFORMATION:
APPLICANT: Griffech, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
TITLE OF INVENTION: Allergenic Proteins And Peptides From
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 514 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-023-134

Query Match 32.9%; Score 787; DB 3; Length 514;
Best Local Similarity 40.5%; Pred. No. 1,3e-62;

Matches 168; Conservative 73; Mismatches 160; Indels 14; Gaps 7;

QY 45 AHDPQAVL---SYLSNIESNNNNIDKVD--KNGIKVIVNLSFGAKGDKGTVDNIAFQA 98
DB 22 AEDQSAQIMLDSYVEKYLRSNRSRLKRVHSRHDANINFERKYGAVGKGDCTEAFSTA 81
QY 99 WNAAGSRTVPQVYVKNKNYLLKQITTSGPCRSSISVKIFGSLBASAKISDYKDRRLMI 158
DB 82 WQACANBS-ANMLVPGSKKFFVNNLFFNGPCOPHFTFVGGIIAAYQNPASWKNRIML 140
QY 159 AFDVONTLVVGGGTINGNGVWMPSSCK-INKSLPC--RAPATLTTWNCNKKVNNLK 215
DB 141 QFAXLIGFTLMGKVIDGGKQWMAQCKWNGREICNDRPRAIKFDFSTGLIIQGLK 200
QY 216 SQAQOIHFKESCTNNVVASNLMINASAKSPNTDGVHVSNTQYQISDTIIGTGDICSI 275
DB 201 LMSBPEFHLVFNCGGVKIGISITAPRDSPTDIDIFASKNPHLQKNTIGTDDCVAI 260

QY 276 VGSQNVQATNITGPGHGISTIGSGNSEAYSVNTVNEAKIIGAENGRIKTKQGS 335

DB 261 GTGSSNIYIEDLIGPBGHGISISLGRNSRAEVSYVHNGAKFIDTONGLRITKTKQGS 320

QY 336 QGASNIKFLNEMQDVKYPPIIIDQNYCDRVPCIQFSAVQVNVYENIKGTSATKVAI 395

DB 321 GMASHIYEVNEMINSNPILINQFCTASASACQNSRNVQIDQVYKRIKTSATMAAI 380

QY 396 KFDSTNPOBGIEMENINVGSGKPSKSEATCNVFN--AEHTPHCTSLKIS 448

DB 381 QKCSDSMPCKDKIKLSIKLTSRK--IASCLINDNANGVFSGHVLPACKNLSPS 433

RESULT 7

5453566-2
Patent No. 5453566
APPLICANT: SHENMAKER, CHRISTINE K.;KRIDL, JEAN C.;HIATT,
WILLIAM R.;KNAUF, VIC
TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
IN PLANT/CELLS
NUMBER OF SEQUENCES: 2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/750,505
FILING DATE: 27-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 240,408
FILING DATE: 30-AUG-1988
APPLICATION NUMBER: 920,574
FILING DATE: 17-OCT-1986
APPLICATION NUMBER: 845,676
FILING DATE: 28-MAR-1986
SEQ ID NO:2
LENGTH: 70
5453566-2

Query Match 15.3%; Score 365; DB 6; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.1e-25;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GIKYINVLSPGAKGDKTYDNIATFEQAMNACSSRTPVQVVPKKNYLLKQITTSFGPCR 131

DB 1 GIKYINVLSPGAKGDKTYDNIATFEQAMNACSSRTPVQVVPKKNYLLKQITTSFGPCR 60

QY 132 SSISVKIFGS 141

DB 61 SSISVKIFGS 70

RESULT 8

5453566-2
Patent No. 5453566
APPLICANT: SHENMAKER, CHRISTINE K.;KRIDL, JEAN C.;HIATT,
WILLIAM R.;KNAUF, VIC
TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
IN PLANT/CELLS
NUMBER OF SEQUENCES: 2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/750,505
FILING DATE: 27-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 240,408
FILING DATE: 30-AUG-1988
APPLICATION NUMBER: 920,574
FILING DATE: 17-OCT-1986
APPLICATION NUMBER: 845,676
FILING DATE: 28-MAR-1986
SEQ ID NO:2
LENGTH: 70
5453566-2

Query Match 15.3%; Score 365; DB 6; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.1e-25;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GIKIVNLSFGAKGDKTYDNINAFEOAMNEACSSRTVPVQVFNKKNYLLKQITTFSGPCR 131
DB 1 GIKIVNLSFGAKGDKTYDNINAFEOAMNEACSSRTVPVQVFNKKNYLLKQITTFSGPCR 60
QY 132 SSISVKIFGS 141
DB 61 SSISVKIFGS 70

RESULT 9
US-08-467-023-189
Sequence 189, Application US/08467023
Patent No. 6090386
GENERAL INFORMATION:
APPLICANT: Griffee, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Shu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-467-023-189

Query Match 13.3%; Score 319; DB 3; Length 127;
Best Local Similarity 56.0%; Pred. No. 3.8e-21;
Matches 65; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

QY 293 HGISISLSGNEBAYVNTVNEAKIIGANGVRITKTQSGSGQASNIKFLVNEQDYK 352
DB 8 HESISISLGRNBRASVYHVAGAFIDPQNGRLRIKTQSGSGMSHIIYENVEVINS 67
QY 353 YPIIDONTCDRVEPCIQFSASAVQVKNVYENIKGTISATKVAIKFDCSTNPFCEGI 408
DB 68 NPIILNQFCTASACONRSASVQIDVTVYKNIRGTSATMAAIQLKCSDSMPCKDI 123

RESULT 10
US-09-107-532A-6652
Sequence 6652, Application US/09107532A
Patent No. 6563275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6652:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...451
SEQUENCE DESCRIPTION: SEQ ID NO: 6652:
US-09-107-532A-6652

Query Match 12.5%; Score 299; DB 4; Length 451;
Best Local Similarity 27.2%; Pred. No. 1.6e-18;
Matches 99; Conservative 65; Mismatches 132; Indels 68; Gaps 15;

QY 73 IXYINLSFGAKGDKTYDNINAFEOAMNEACSSRTVPVQVFNKKNYLLKQITTFSGPCR 132
DB 5 IMYDILTFGASID--ELNTEAIIQQAIDMAASDGGT-VVVPAGE-----FLTGALFL 54
QY 133 SISVKIFGSLBASKIS-DYKD-----RLW---IAFSDVONLVVGGGTING 176
DB 55 KSNVEIHLASGAVLKSSDPKDYPVVHSRVEGHRKRVASCIYAQVENVISVTGFTLDG 114
QY 177 NGQVWV-----PSSCKINKSLPCRDAPYALTFWNCNKLKVNVLKSKNAQQIHIKESCT 230
DB 115 NGKKWHTFRNEPD---NLAYP---RPLKMSPHNCHRIIVDKIKIQSPSWTINPILCS 167
QY 231 NVVASLMLNNAKASNTGVAVSNTQYIQISDTIIGTDDDCISYSGS-----QN 281
DB 168 NATFDNLTLINPADSNTDGIPESCKNVRISNCHLIDVDDCIAIAGTEDTYERIACEN 227

Boot LocalSimilarity 25.1%: Pred. No.1,8e-18		
Matches 107:	Conservative 73:	Mismatches 166: Indels 8: Gaps 16
QY	74 KVINLVSFGAKGDKTYDNINIAFEQAMNEACSSRTFVQFVPRKNKYLL--KQITFSGPC	130
DB	48 KTCVHRSHGSDTDDSDYILSALNQ-----CNHGKRVF--DEKKEYIIGTLNMTF----	96
QY	131 RSSLSVAKIFGSL-----EASSKSIDXYKDRRLMIAFDSVQNLVGGGGGTINNGQVW	181
DB	97 LKNIDLEVLGTLTLNTDIDYQWANSFKQFQWATTFPQGLG--EDVNMGGGSGTINNGQVW	155
QY	182 WPSSCKINKSPCRDAPALTFFWNCNKLKVNLLSKSNAQOIHKEESCTNVVASNLMTNA	241
DB	156 YDLVAEDDLI----RPLIMGIGLNGGTIGPLKLRYPQYHFPANSSNVLPDGDIDSG	211
QY	242 SAKS---PNTDGVHVSNTQYIQISDITIGTGDCDISVSGSQNVQATNITGCPGHGISI	297
DB	212 YSKSDNEAKNTGDWMDTVRSNNIIVLONSVINNGDDCVSPKFNSTNLVQNLHCNGSHGISV	271
QY	298 GSLGSGNSE-----AYVSNVTYVNEAKIIGANGV-----RIKTMQ-----	332
DB	272 GSLQYQKDEVDIVENVVYVYINISMPNASCLENFHHIIDLTLTLODMARIKVPGRPSALS	331
QY	333 -----GSSGQASNIKFLNVEQVQYKPIIIDQNYCDR--VEBCIOQFSAVQKVVYENI	385
DB	332 ADLGGGGSSGVKNITVDTALIDNDAIMEIQCGQKNTLLCNEYPPSLITSDVHIKNF	391
QY	386 KGISA-----TKVAIKFDCSTNFPCEGIMENINLVGSGSKPSEATCKNVHFNNNAEH	437
DB	392 RGTTSGESEDPYVGIV-----CSSPDTCSDIYTSININTVSPDG--TNDPVCNVV---DESL	442
QY	438 VTHPCTS 444	
DB	443 LSVNCTA 449	

```

US-09-787-583-2
RESULT 13
; Sequence 2, Application US/09787583
; Patent No. 6602696
; GENERAL INFORMATION:
; APPLICANT: DSM N.V.
; APPLICANT: Franse, Maartje
; APPLICANT: Grasselin, Catherine
; APPLICANT: Herweijer, Margareta
; APPLICANT: Meuwissen, Petrus
; APPLICANT: Ooijen, Albert
; APPLICANT: Vorrigen, Alphons
; TITLE OF INVENTION: ASPERGILLUS TUBIGENSIS POLYGALACTURONASE
; FILE REFERENCE: 24615-20144.00
; CURRENT APPLICATION NUMBER: US/09/787,583
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: EP 99200481.2
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: EP 98203171.8
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 362
; TYPE: PRF
; ORGANISM: Aspergillus tubingensis
; US-09-787-583-2

```

	Query Match	12.3%	Score 293.5;	DB 4;	length 362;
	Best Local Similarity	28.3%;	Pred. No.	3.6e-18;	
	Matches	80; Conservative	59;	Mismatches	91; Indels 53; Gaps 13
Qy	171	GCTINGNCQVW-----PSSCKINRSLPCRDATATLTFFNNCKNLKYNLKSXA	219		
Db	102	GAKINCDDARMMWDGSGNGCKTKPFQVHK-----LDSEJITELKYNNPVQC-	150		
Yy	220	QQIHKEFSCTNVVASNLMIN-----ASAKSPNTDGHVSNTQYIQISDTIIGTGDD	271		

Db 151 -----PSILADHITIDVYIINDSAGSKGHNTAFDIGSGTYITIDGATYYNQDD 200
 QY 272 CISIVSGSONVQATNITCTGPGHGHSIGSLGSGNSEAYVSNVTUNEAKIIGAENVRIKWTW 331
 Db 201 CLAINSG-EHITFTFNGYCDGGHGHSIGSIG-GRSDNTYNDVYITISKVLNSQNGVRIKWT 258
 QY 332 QGGSGQASNIKFLVEMQDV-KYPIIIDONYCDRAVEPCIQOFSAYOVKNVYENIKG--- 387
 Db 259 YGKGTGYENVAFFEDITLSDISKYGLVQBQY-ENSGPCTPTNGVYVEDITFAKVTGSYK 317
 QY 388 TSATKVAIKFDCSTNPFCEGIIMENINLVGSGKPESEATCKNV 430
 Db 318 SSGTDIYIL--CGSG-SCSNMTWGSVDYTG--GKKS-SKCKNV 354

```

RESULT 14
US-09-542-767A-1
: Sequence 1, Application US/09542767A
: Patent No. 6296671
: GENERAL INFORMATION:
: APPLICANT: Schultein, Martin
: APPLICANT: Kristensen, Henrik
: TITLE OF INVENTION: An Enzymatic Treatment Method
: FILE REFERENCE: 5871, 204-US
: CURRENT APPLICATION NUMBER: US/09/542,767A
: CURRENT FILING DATE: 2000-04-04
: PRIOR APPLICATION NUMBER: PA 1999 00390
: PRIOR FILING DATE: 1999-03-22
: PRIOR APPLICATION NUMBER: US 60/125,884
: PRIOR FILING DATE: 1999-03-24
: PRIOR APPLICATION NUMBER: PCT/DK00/00136
: PRIOR FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 360
: TYPE: prt
: ORGANISM: Trichisporon penicillatum
: US-09-542-767A-1

```

	Query Match	10.8%;	Score 257.5;	DB 3;	Length 360;		
	Best Local Similarity	15.5%;	Pred. No. 6.5e-15;				
	Matches	72;	Conservative	54;	Mismatches	105;	
				Indels	51;	Gaps	10
Qy	171	GGTNGNGQVWPPSSCKTKNSL-PRDAPLTALTFWNCCKLKNNLSKMAQOIHIFESC	229				
Db	104	GSVIDEGEARWMD-----NKGANGCKVPRLEYAHNLNDSHNGLIHNTFPVGFESID-	157				
Qy	230	TNVAASNLIMINASAKSP-----NTDGVHNSNTQYIOISPTIIGTGDPCISIVSGSQWQVAT	285				
Db	158	KNLIIDGVRIODNSDGTGCAFNTDPAFVDSQSYNVATIQAMVYHNODDCLAINQO-ELIHFL	216				
Qy	286	NITGCPGHGISIGLSGNSSEAYVSNVYNEAKTIIGAENGVAIKTWGGSGGASNTKFLN	345				
Db	217	NGYCYGGGGLSIGSVGGGN---VVDVVIABQSIIINSQGVAKIKTKSGGTGEVGRGITYRN	273				
Qy	346	VEMODV-KYPIIITQNYCDRVEPCIQDQPSAVQVKNVVENIKTSA-----	390				
Db	274	IFLGGITDYGILVQDDYNNPGHAT---NSIKIHITFPMVHGATQHGFINIAFCGDS	329				
Qy	391	-----TKVAIKPDCSTNFPCEGIMINILVSGSGKPEATK	428				
Db	330	CYDWTMEVKITHGADYKQNV-----PSSASQ	358				

```

RESULT 15
US-08-061-062A-6
; Sequence 6, Application US/08061062A
; Parent No. 5550045
;
GENERAL INFORMATION:
;
APPLICANT: MUSTERS, WOUTER
;
APPLICANT: STAM, HEIN
;

```

us-10-691-374-2.rai

```

1  APPLICANT: SUYKERBUIK, MARIA E.
2  APPLICANT: VISSER, JACOB
3  APPLICANT: VERBAKEL, Johannes M.
4  TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
5  TITLE OF INVENTION: ENCODING A RIPENING FORM OF A POLYPEPTIDE HAVING
6  TITLE OF INVENTION: RHANNOALACTURONASE ACTIVITY
7  NUMBER OF SEQUENCES: 16
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: CUSHMAN DARBY & CUSHMAN
10 STREET: 1100 NEW YORK AVENUE, N.W.
11 CITY: WASHINGTON, D.C.
12 COUNTRY: U.S.A.
13 ZIP: 20005-3918
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Disk
17
18 COMPUTER: IBM PC compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: PatentIn Release #1.0, Version #1.25
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/061,062A
23 FILING DATE: 14 MAY 1993
24 CLASSIFICATION: 435
25 ATTORNEY/AGENT INFORMATION:
26 NAME: KOKULIS, PAUL N.
27 REGISTRATION NUMBER: 16773
28 REFERENCE/DOCKET NUMBER: 202390/R 7262 (V)
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (202) 861-3000
31 TELEFAX: (202) 822-0944
32 TELEEX: 6714627 CUSH
33 INFORMATION FOR SEQ ID NO: 6:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 440 amino acids
36 TYPE: amino acid
37 TOPOLOGY: linear
38
39 MOLECULE TYPE: protein
40
41 IS-08-061-062A-6

```

QY 70 KNGIKVNVLSFGAKDDGKTYDNIAEQWANEACSSRTVPQVPEVPKNKVVILKOITFSGP 129
 Db 33 KAKTKCNILISYGAVDNSTDVGPAITSAW-AAKCSGGLV-YIISGVYALNTWVTLTGG 89
 QY 130 CRESSISV--IFGSLSSAKSIDPKXRRLIMAFDSVQULVVGGGGTINGNQVWPPSSCK 187
 Db 90 SAIATOLDGIIYRTGSAAGNMIAVDTTDFELFSSTSK-----GAVGQFYVYHABGT 143
 QY 188 INKSLPCRPAPALIFWNCNKILKVNLSKNAOQIHIKFECSCTWVVASNLMTNANSASP 247
 Db 144 GARILRLTV-----THRSVADVILYDAPAFHFTMDCSIGSEVYNAIKRGNEG-G 193
 QY 248 TDGVHVSNTQYLOISDTLIIGTGDDCISIVSSQVNVQATNITTCGGHGHSIGSLGSGNSEA 307
 Db 194 LDGIDVWGSN-IWHDDEVTKRDECVYKSPANNILVESIYCWMSGGCGAMGSLG---ADT 249
 QY 308 YVSNVTVEAKIIGAENGVRIKTQWGSQGAASNIKFLVNMQVKKPIIIDQVYCDRVEP 367
 Db 250 DVTIDIVRNVYTWSSNMOMYIKS-NGSGSVYVNVLENTFGHNAVASLIDIGWSSMTAV 308
 QY 368 CIOQFSAVQVKNVYENIKGT-----SATKAIAKDCSTNFPCEGIIMENINLVGSGSKPSE 424
 Db 309 A---GDGVQNLNNTIVKMKGTGANATRPPIRVVCSDTAPCTDLTLEDIAIWTBSSSEL 365
 QY 425 ATCKNVH 431
 Db 366 YLCRSAY 372

Search completed: March 24, 2005, 07:06:49
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 06:59:04 ; Search time 169 Seconds
(without alignments)
1045.855 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MWIGRNSILLILIIIFASIS.....VTPHCTSLSEIDELVNY 457

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	457	1	AAP80299 AAP80299 Polygalact
2	2390	100.0	457	1	AAP94619 AAP94619 Polygalact
3	2390	100.0	457	2	AAR32107 AAR32107 Polygalact
4	2390	100.0	457	4	AAB84338 AAB84338 Tomato po
5	2338.5	97.8	456	6	ABR43936 ABR43936 Tomato en
6	1185	49.6	438	5	ABR92243 ABR92243 Herbicida
7	1127	47.2	431	2	AAW98178 AAW98178 Anthor-sp
8	1127	47.2	431	3	AAW41324 AAW41324 Arabidops
9	1127	47.2	431	5	ABR92704 ABR92704 Herbicida
10	1127	47.2	431	3	AAW41323 AAW41323 Arabidops
11	1119.5	46.8	415	3	AAW41325 AAW41325 Arabidops
12	1115.5	46.2	426	5	ABR92075 ABR92075 Herbicida
13	1103	46.2	433	2	AAW04268 AAW04268 Ollseed r
14	1103	46.2	433	2	AAW42649 AAW42649 Brassica
15	1056.5	44.2	404	2	AAR48674 AAR48674 Putative
16	925	38.7	452	4	AAE00422 AAE00422 Tomato po
17	888.5	37.2	441	5	ABR91673 ABR91673 Herbicida
18	845	35.4	436	3	AAW52347 AAW52347 Arabidops
19	845	35.4	459	5	AAW52346 AAW52346 Arabidops
20	845	35.4	459	5	ABR91108 ABR91108 Herbicida
21	840.5	35.2	468	5	ABR91537 ABR91537 Herbicida
22	796.5	33.3	514	2	AAW42122 AAW42122 Japanese
23	791.5	33.1	514	2	AAW04346 AAW04346 Chamaecyp
24	787.5	32.9	434	5	ABR91377 ABR91377 Herbicida
25	787.5	32.9	514	2	AAR74333 AAR74333 Japanese

26	787.5	32.9	514	2	AAR81586 AAR81586 Cedar pol
27	787.5	32.9	514	2	AAY25667 AAY25667 Japanese
28	787.5	32.9	514	7	ADC34915 ADC34915 Cedar all
29	787	32.9	514	2	AAR53690 AAR53690 Japanese
30	787	32.9	514	2	AAR69792 AAR69792 Japonicum
31	787	32.9	514	2	AAR93599 AAR93599 Japan ced
32	787	32.9	514	2	AAY25666 AAY25666 Japanese
33	787	32.9	514	7	ADC34914 ADC34914 Cedar all
34	767	32.1	460	2	AAR69791 AAR69791 Japonicum
35	764	32.0	453	4	AAM51693 AAM51693 Juniperus
36	764	32.0	507	4	AAM51691 AAM51691 Juniperus
37	741.5	31.0	491	5	ABR90812 ABR90812 Herbicida
38	728	30.5	405	5	ABR92113 ABR92113 Herbicida
39	724.5	30.3	398	5	AAE20568 AAE20568 Pear poly
40	718	30.0	435	5	ABR93472 ABR93472 Herbicida
41	718	30.0	482	3	AAW05118 AAW05118 Arabidops
42	715	29.9	392	5	ABR92116 ABR92116 Herbicida
43	709.5	29.7	515	5	ABR90945 ABR90945 Herbicida
44	708.5	29.6	401	3	AAW41452 AAW41452 Arabidops
45	708.5	29.6	422	3	AAW41451 AAW41451 Arabidops

ALIGNMENTS

RESULT 1	AAP80299	standard; protein; 457 AA.
ID	AAP80299	
XX		
AC	AAP80299;	
XX		
DT	25-MAR-2003 (revised)	
DT	20-NOV-1990 (first entry)	
XX		
DE	Polygalacturonase.	
XX		
KW	Fruit ripening; polygalacturonase; pectin esterase.	
XX		
OS	Unidentified.	
XX		
PN	EP271988-A.	
XX		
PD	22-JUN-1988.	
XX		
PF	06-NOV-1987; 87EP-00309853.	
XX		
PR	11-NOV-1986; 86GB-00026879.	
XX		
PA	(ICIL) IMPERIAL CHEM IND PLC.	
PA	(ZENNE) ZENNECA LTD.	
PI	Bridges IG, Schuch WW, Grierson D;	
XX		
DR	WPI; 1988-169271/25.	
DR	N-PSDB; AAN80487.	
XX		
PT	Recombinant DNA comprising promoter and terminator sequences - useful in	
PT	plants for altering ripening properties esp. in tomatoes.	
XX		
PS	Disclosure; Page 7; 22pp; English.	
XX		
CC	This polygalacturonase (PG) is encoded by plasmid clone PROM6 which is	
CC	used to produce antisense mRNA (with an inverted sequence to that of PG	
CC	mRNA) which is inserted into a vector used to transform plants which	
CC	thereafter have altered ripening properties. The inverted sequence and	
CC	the PG mRNA form a double-stranded structure which inhibits ex- pression	
CC	of the PG mRNA. See also AAN80488. (Updated on 25-MAR-2003 to correct PA	
CC	field.)	
XX		
SQ	Sequence 457 AA;	
Query Match	100.0%; Score 2390; DB 1; Length 457;	
Best Local Similarity	100.0%; Pred. No. 1.1e-193;	

Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
Db 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIJAEQAMNEACSSRTVQVFPKKNXYL 120
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIJAEQAMNEACSSRTVQVFPKKNXYL 120
QY 121 LKQITFSGPCRSISVKIFGSLFASASKISDYKDRRLMTAFDSVONLVVGGGTTNGNGOV 180
Db 121 LKQITFSGPCRSISVKIFGSLFASASKISDYKDRRLMTAFDSVONLVVGGGTTNGNGOV 180
QY 181 WMPSSCKIKKSLPCRDAPFALTFFMNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
Db 181 WMPSSCKIKKSLPCRDAPFALTFFMNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIOISDPTIIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
Db 241 ASAKSPNTDGVHVSNTQYIOISDPTIIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
QY 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKWOGSGGASNIKFLANEMODVKYPIIIIDQN 360
Db 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKWOGSGGASNIKFLANEMODVKYPIIIIDQN 360
QY 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSEG 420
Db 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSEG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457

```

RESULT 2

AAP94619

ID AAP94619 standard; protein; 457 AA.

XX AAP94619;

AC

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC peptide which may be joined to heterologous peptides directing them to
 CC cell wall. (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 457 AA;

Query Match
 Best Local Similarity 100.0%; Score 2390; DB 1; Length 457;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
Db 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIJAEQAMNEACSSRTVQVFPKKNXYL 120
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIJAEQAMNEACSSRTVQVFPKKNXYL 120
QY 121 LKQITFSGPCRSISVKIFGSLFASASKISDYKDRRLMTAFDSVONLVVGGGTTNGNGOV 180
Db 121 LKQITFSGPCRSISVKIFGSLFASASKISDYKDRRLMTAFDSVONLVVGGGTTNGNGOV 180
QY 181 WMPSSCKIKKSLPCRDAPFALTFFMNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
Db 181 WMPSSCKIKKSLPCRDAPFALTFFMNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIOISDPTIIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
Db 241 ASAKSPNTDGVHVSNTQYIOISDPTIIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
QY 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKWOGSGGASNIKFLANEMODVKYPIIIIDQN 360
Db 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKWOGSGGASNIKFLANEMODVKYPIIIIDQN 360
QY 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSEG 420
Db 361 YCDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSEG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457

```

RESULT 3

AAR32107

ID AAR32107 standard; protein; 457 AA.

XX AAR32107;

AC

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

polygalacturonase; PG; pectin esterase; PG; expression regulation;
 fruit softening enzymes; flowering plants; fruiting plants;
 antiense RNA.

Lycopersicon esculentum.

EP532060-A1.

17-MAR-1993.

06-NOV-1987; 92EP-00117411.

11-NOV-1986; 86GB-00026879.

(ICIL) IMPERIAL CHEM IND PLC.

(ZENNE) ZENNECA LTD.

Bridges IG, Grierson D, Schuch WW;

CDNA of gene may act as a probe to the genomic sequence, provides a means
 of modulating the production of PG and acts as a source of the transit

DR WPI; 1993-087084/11.
 DR N-PSDB; AAQ38415.
 PT Recombinant DNA for flowering and fruiting plants e.g. tomatoes ripening
 PT control - comprises base sequence for transcription contg. inverted
 PT sequence of bases complementary to bases in anti sense ribonucleic acid
 PT encoding softening enzymes, or gene expression regulation.
 XX
 PS Example 12; Fig 1; 20pp; English.
 CC This is the sequence of polygalacturonase from clone pTOM6. The clone was
 CC used to isolate the PG promoter in the construction of a vector encoding
 CC antisense RNA to the PG cDNA and PG gene. This would be useful to
 CC regulate the expression of the fruit softening enzymes in flowering and
 CC fruiting plants. Such antisense RNA would delay fruit softening. (Updated
 CC on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct
 CC PE field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-
 CC AVG-2003 to correct OS field.)
 CC
 XX
 SQ Sequence 457 AA;

Query Match 100.0%; Score 2390; DB 2; Length 457;
 Best Local Similarity 100.0%; Pred. No. 1.1e-193;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVIORNSILLIIIFASSISTCRSNVIDNLFQOVYDNIIEQFADHFOAYLSYLSKNIE 60
 DB 1 MVIORNSILLIIIFASSISTCRSNVIDNLFQOVYDNIIEQFADHFOAYLSYLSKNIE 60
 QY 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120
 DB 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120
 QY 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180
 DB 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180
 QY 181 WMPSSCKINKSLPCRDAPFALTFFWNCNKLKVNNLKSKNAQOIHKFESCTNVVASNLMIN 240
 DB 181 WMPSSCKINKSLPCRDAPFALTFFWNCNKLKVNNLKSKNAQOIHKFESCTNVVASNLMIN 240
 QY 241 ASAKSPNTDGVHNSNTQYIOISPTIIGTGDDCISIVSGSONVQATNITCGPGHGISISGL 300
 DB 241 ASAKSPNTDGVHNSNTQYIOISPTIIGTGDDCISIVSGSONVQATNITCGPGHGISISGL 300
 QY 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGSGQASNIKFLANVMQDVKYPPIIIDON 360
 DB 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGSGQASNIKFLANVMQDVKYPPIIIDON 360
 QY 361 YCDRVEPCIOQFSAVOVKVNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420
 DB 361 YCDRVEPCIOQFSAVOVKVNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420
 QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457
 DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457

RESULT 4
 AAB48338
 ID AAB48338 standard; protein; 457 AA.
 XX
 AC AAB48338;
 XX
 DT 20-APR-2001 (first entry)
 DE Tomato polygalacturonase (PG) enzyme.
 XX
 KM Pectin; pectin methylase; PME; polygalacturonase; PG; tomato; pTOM6;
 KM food product; yogurt; milk; fruit juice; whey drink; de-esterification.
 XX
 OS Lycopersicon esculentum.
 XX

PN WO200078982-A1.
 XX
 PD 28-DEC-2000.
 XX
 XX 15-JUN-2000; 2000WO-1B000869.
 XX
 XX 17-JUN-1999; 99GB-00014209.
 XX
 PA (DANI-) DANISCO AS.
 XX
 PI Christensen TME, Kreiberg JD;
 XX
 DR WPI; 2001-091573/10.
 DR N-PSDB; AAC84653.
 XX
 PT Modifying pectin, for foodstuffs preparation, involves transforming host
 PT having pectin methylase (PME) and polygalacturonase (PG) activity by
 PT silencing PG activity, to increase PME to PG ratio.
 XX
 PS Disclosure; Fig 1; 78pp; English.

CC The invention provides a new method for modifying pectin that involves
 CC providing a host having pectin methylase (PME) activity and
 CC polygalacturonase (PG) activity, transforming the host by silencing PG
 CC activity to provide an increased PME to PG ratio, preparing a PME extract
 CC from the transformed host, and using the PME extract to modify pectin. A
 CC PME modified pectin is useful for foodstuffs preparation, and to impart
 CC an increased functionality to food products such as yogurt, milk/fruit
 CC juice and whey drinks. PME is useful to reduce the number of ester groups
 CC in a pectin in a block-wise manner, and to de-esterify two or more
 CC adjacent galacturonic acid residue of a pectin on at least substantially
 CC all of the pectin chains. The present sequence represents a PG enzyme
 CC encoded by a pTOM6 cDNA
 CC
 XX
 SQ Sequence 457 AA;

Query Match 100.0%; Score 2390; DB 4; Length 457;
 Best Local Similarity 100.0%; Pred. No. 1.1e-193;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVIORNSILLIIIFASSISTCRSNVIDNLFQOVYDNIIEQFADHFOAYLSYLSKNIE 60
 DB 1 MVIORNSILLIIIFASSISTCRSNVIDNLFQOVYDNIIEQFADHFOAYLSYLSKNIE 60
 QY 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120
 DB 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120
 QY 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180
 DB 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180
 QY 181 WMPSSCKINKSLPCRDAPFALTFFWNCNKLKVNNLKSKNAQOIHKFESCTNVVASNLMIN 240
 DB 181 WMPSSCKINKSLPCRDAPFALTFFWNCNKLKVNNLKSKNAQOIHKFESCTNVVASNLMIN 240
 QY 241 ASAKSPNTDGVHNSNTQYIOISPTIIGTGDDCISIVSGSONVQATNITCGPGHGISISGL 300
 DB 241 ASAKSPNTDGVHNSNTQYIOISPTIIGTGDDCISIVSGSONVQATNITCGPGHGISISGL 300
 QY 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGSGQASNIKFLANVMQDVKYPPIIIDON 360
 DB 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGSGQASNIKFLANVMQDVKYPPIIIDON 360
 QY 361 YCDRVEPCIOQFSAVOVKVNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420
 DB 361 YCDRVEPCIOQFSAVOVKVNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420
 QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457
 DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLVNY 457

RESULT 5

ABR43936
ID ABR43936 standard; protein; 456 AA.
XX
AC ABR43936;
XX
DT 11-AUG-2003 (first entry)
XX
DE Tomato endopolysaccharonase enzyme.
XX
DE Plant; pectin transferase; xyloglucan; polylysine; heparin;
XX
KW anticoagulant; endopolysaccharonase; enzyme; tomato.
XX
OS Lycopersicon esculentum.
XX
PN WO2003017950-A2.
XX
PD 06-MAR-2003.
XX
PF 03-SEP-2002; 2002WO-US028066.
XX
PR 31-AUG-2001; 2001US-0316777P.
XX
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
PI Albersheim P, Djelineo-Albersheim I, Davill A;
XX
DR WPI; 2003-421044/39.
XX
PT Formation of ester or amide bond between ester and alcohol or amine
PT Involves treating the ester or its acid or salt with plant pectin
XX transferase synthase in the presence of alcohol or amine.
PS Disclosure; Fig 18; 67pp; English.
XX
CC The invention relates to forming an ester or amide bond between monomeric
CC or polymeric ester or its acid or salt and monomeric or polymeric alcohol
CC or amine by treating the ester or its acid or salt with a plant pectin
CC transferase synthase in the presence of alcohol or amine under conditions
CC to form ester or amide bond. The method is useful for forming an ester or
CC amide bond between monomeric or polymeric ester or its acid or salt (e.g.
CC homogalacturonan) and monomeric or polymeric alcohol or amine. It is
CC useful for producing pectin-based polymers e.g., xyloglucan or D- or L-
CC polylysine useful for the slow release of compounds in the body e.g.,
CC heparin with anticoagulant or other pharmaceutical properties. The
CC present sequence represents an endopolysaccharonase enzyme from tomato
XX
SQ Sequence 456 AA;

Query Match 97.8%; Score 2338.5; DB 6; Length 456;
Best Local Similarity 98.9%; Pred. No. 2.6e-189;
Matches 452; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 MWIQNSLILIIIFASSISTCRSNVDNLFKQVVDNLLEQFAHDPQAVLSYLSKNIE 60
DB 1 MWIQNSLILIIIFASSISTCRSNVDNLFKQVVDNLLEQFAHDPQAVLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKIVIVLSFGAKGDKTYDINAEQAWNEACSSRTVPQVVKKNKYL 120
DB 61 SNNNIDKVDKNGIKIVIVLSFGAKGDKTYDINAEQAWNEACSSRTVPQVVKKNKYL 120
QY 121 LKQITSGPCRSISYKIFGSLASAKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGOV 180
DB 121 LKQITSGPCRSISYKIFGSLASAKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGOV 180
QY 181 WPPSSCKIKSLPCRDAPALTFFMCKLKNKLNKSNAAQOIHIFKESCTNVVANSMLMN 240
DB 181 WPPSSCKIKSLPCRDAPALTFFMCKLKNKLNKSNAAQOIHIFKESCTNVVANSMLMN 240
QY 241 ASAKSPNDGAVVNTQYTIQISDTTIGTGDPCISIVSGSONVOATMTGCPGHGISTSL 300
DB 241 ASAKSPNDGAVVNTQYTIQISDTTIGTGDPCISIVSGSONVOATMTGCPGHGISTSL 300

QY 301 GSGNSEAYVSNVTVAEAKIIGAENGVRIKTWQGGSGQASNIKPLANVMODVKYPIIIDON 360
DB 301 GSGNSEAYVSNVTVAEAKIIGAENGVRIKTWQGGSGQASNIKPLANVMODVKYPIIIDON 360
QY 361 YCDRVPEPCIQFSAVQVKNVYVENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSG 420
DB 361 YCDRVPEPCIQFSAVQVKNVYVENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSG 420
QY 421 KPSEATCKVHFNNAEHVTPHCTSLIESEDEALLNY 457
DB 421 KPSEATCKVHFNNAEHVTPHCTSLIESEDEALL-NY 456

RESULT 6

ABB92243
ID ABB92243 standard; protein; 438 AA.
XX
AC ABB92243;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 1454.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidner M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
PS Claim 5; SEQ ID NO 1454; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC herbicides
XX
SQ Sequence 438 AA;

Query Match 49.6%; Score 1185; DB 5; Length 438;
Best Local Similarity 51.9%; Pred. No. 1.6e-91;
Matches 230; Conservative 76; Mismatches 121; Indels 16; Gaps 5;
QY 8 ILLIIIFASSISTCRSNV--IDNLFKO---VYDNLLEQFAHDPQAVLSYLSKNIES 61
DB 5 ILLISFLITLPLSCFSSYPFNHRDLFMSNNVYLETNNKHQHGN-----TRNSHL 56
QY 62 NNNIDKVDKNGIKIVIVLSFGAKGDKTYDINAEQAWNEACSSRTVPQVVKKNKYL 121
DB 57 KNRHGYAPRRSPRFNNVTFGAAGAND-DSKAFMFAWEACSSGTGIYIVAPKRDVIML 115
QY 122 KQITSGPCRSISYKIFGSLASAKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGOV 181

```

Db      116 KATTFGCPCKSLIFLYGRIEAMENPSDYKERHWHVFENNANLREVGGRIDGNGHW 175
Qy      182 WPSSCKINSLPCRDAPALTPMNCNKLKNNLSKNAQOIHKFESCTNVASNMINA 241
Db      176 WPSSCKINPOLPLGATPAVATFVECNRLRNSRLNLAQOMHLTFDDCKRVKLNLMVTS 235
Qy      242 SAKSPNTDGVHVNTOYIOISDTIIGTGDPCISIVSGSQVNAQNTNITCGPHGISISGLG 301
Db      236 PADSPNTDGHVSTGTONIILQDSIVRTGDDCISIVSGSENVRAITGTCGPHGISISGLG 295
Qy      302 SGNSSEAVSNVTNEAKIIGANGVRIKTIQSGSGAASNIKFLVEMQDVYPIIIDQNY 361
Db      296 EDNSEAVSNVNVVVKATLTIGTTNGVRIKTIQSGHGMANKNIIPODIIMKNTNIIINQDY 355
Qy      362 CDRVEPCIQFSAVQVKNVYENIKGSATKVAIKFDCSTNPFCEGIMENINLVGESGK 421
Db      356 CDRVEACPCKSAVQVSNVLYKNIOGTSRPIAVKFCVSKNIPCRGISMQNVKLVDTQTO 415
Qy      422 P-SEATCKNVHFNNAEHVTPHCT 443
Db      416 DVSKASCSNVKLDTRGNVSPFCT 438

RESULT 7
AAW98178 standard; protein; 431 AA.
XX      AAW98178;
XX      DT      05-JUL-1999 (first entry)
XX      DE      Anther-specific E5J2A protein.
XX      KW      E5J2A gene; promoter; pollen; anther dehiscence; male sterile;
XX      KW      transgenic plant.
XX      OS      Arabidopsis thaliana.
XX      PN      W09913089-A1.
XX      PD      18-MAR-1999.
XX      PF      11-SEP-1998; 98WC-GB002752.
XX      PR      11-SEP-1997; 97GB-00019359.
XX      PA      (BIOG-) BIOGENMA UK LTD.
XX      PI      Roberts JA, Paul W, Craze M;
XX      DR      WPI, 1999-254279/21.
XX      DR      N-PSDB; AAX25011.
XX      PT      Generation of male sterile plants by controlling anther dehiscence.
XX      PS      Disclosure; Fig 3; 34pp; English.
XX      CC      This protein is encoded by the E5J2A gene (see AAX25011) of Arabidopsis
XX      CC      thaliana. The invention relates to the use of the E5J2A promoter to
XX      CC      reduce dehiscence and to create male sterile plants for use in hybrid
XX      CC      seed production. The promoter is used to drive expression of a further
XX      CC      nucleic acid sequence that results in prevention or reduction of anther
XX      CC      dehiscence. For example, expression of the Ruase barnase causes cell
XX      CC      ablation, while expression of a plant hormone alters the developmental
XX      CC      fate of a cell. Plants are produced that have phenotypically normal
XX      CC      pollen grains, within phenotypically normal anthers, but in which the
XX      CC      anthers do not dehisce and thus do not release the pollen grains. The
XX      CC      system allows the female to be multiplied with the artificial male
XX      CC      sterility gene in the homozygous state, since the female plant produces
XX      CC      viable pollen. The system is suited to crops which have high seed
XX      CC      multiplication, large amounts of pollen, and/or separate male and female
XX      CC      inflorescences. These factors allow for easy collection of pollen from

```

```

CC      non-dehiscent anthers, facile self-pollination and the minimisation of
CC      the area of plants that have to be self-pollinated manually. Such an
CC      ideal crop is the monocot maize, but the system is also applicable to
CC      e.g. wheat, barley, rice, fodder grass, banana, palm, orchid, tulip,
CC      lily, melon, cucumber, tomato, pepper and willow. It is also useful in
CC      the avoidance or reduction of pollen allergens and may be effective in
CC      the control of asthma caused by pollen release
XX      SO      Sequence 431 AA;
Qy      Query Match      47.2%; Score 1127; DB 2; Length 431;
Qy      Best Local Similarity 49.6%; Pred. No. 1,3e-86;
Qy      Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;
Qy      7 STLLIIIFASSISTCRSNVID-----DNLFKQVONNILEQFAHDFQAVLSYL 55
Db      9 AVFLCVLMLSLCKALSSVNDGTYGHEDSFPSSDSLKLKLNDDVL-----SLISSD 59
Qy      56 SKNIESNNNIDKVDKNGIVNVLSPGAKGDKTYDNIAFEQAMNEACSSRTPVQFVPR 115
Db      60 ETLLEAS-----TVGSNFGAKGDKGTDGTFQAFKAMKKAAGSTNGVTFLEVR 107
Qy      116 NKNYLLKQITFGPCRSISVKKIFGSLBASSKISDYKDRMLAFPSVQNLVVGCG--GT 173
Db      108 GKTYYLTKSTRFRPCKSLRNFOILGTLASATYKSDYKDKNHLILEDVNLSIDGSTGI 167
Qy      174 INGNQVWMPSSCKIKSLPCRDAPALTPMNCNKLKNNLSKNAQOIHKFESCTNVV 233
Db      168 INGNKTMQNSCKIDSKPCCTKAPALTLVNLKLNANLKRKNAQOIOISEKNRVE 227
Qy      234 ASNLMINASAKSPNTDGVHVSNTQYIOISDTIIGTGDPCISIVSGSQVNAQNTNITGPGH 293
Db      228 VSNVEITAPGDSGNTDGHITMTQNIIRVNSNDIGTDDCISIBDGTQNIIFDLTCGPH 287
Qy      294 GISIGSLGSGNSPARYSNVTNEAKIIGANGVRIKTIQSGSGAASNIKFLVEMQDVYK 353
Db      288 GISIGSLGSDNSKAYVSGINVDGAKFESBNGVRIKTIQSGSGTAKIKFQNIIRMNVN 347
Qy      354 PIITDQYCDRVEPCIQFSAVQVKNVYENIKGSATKVAIKFDCSTNPFCEGIMENI 413
Db      348 PIITDQYCDK-DKCDQESAQVQVKNVYKNISGTSATDAIYILNCSBKPCGQIVLENV 406
Qy      414 NLVSGSKPSEATCKNVHFNNAEHVTPHCT 443
Db      407 KIKG-----GTASCKNVANXGCTVSPKCS 431

RESULT 8
AAG41324
ID      AAG41324 standard; protein; 431 AA.
XX      AC      AAG41324;
XX      DT      18-OCT-2000 (first entry)
XX      DE      Arabidopsis thaliana protein fragment SEQ ID NO: 51400.
XX      KW      Protein identification; signal transduction pathway; metabolic pathway;
XX      KW      hybridisation assay; genetic mapping; gene expression control; promoter;
XX      KW      termination sequence.
XX      OS      Arabidopsis thaliana.
XX      PN      EP1033405-A2.
XX      PD      06-SEP-2000.
XX      PF      25-FEB-2000; 2000EP-00301439.
XX      PR      25-FEB-1999; 99US-0121825P.
XX      PR      05-MAR-1999; 99US-0123180P.
XX      PR      09-MAR-1999; 99US-0123548P.
XX      PR      23-MAR-1999; 99US-0125788P.

```

PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129645P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138647P.
PR 14-JUN-1999; 99US-0139112P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 24-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140659P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 02-JUL-1999; 99US-0141842P.
PR 06-JUL-1999; 99US-0142154P.
PR 08-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145192P.
PR 26-JUL-1999; 99US-014524P.
PR 27-JUL-1999; 99US-014524P.
PR 27-JUL-1999; 99US-0145913P.
PR 28-JUL-1999; 99US-0145918P.
PR 02-AUG-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146388P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 07-SEP-1999; 99US-0153070P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0155659P.
PR 29-SEP-1999; 99US-0156458P.
PR 04-OCT-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157177P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.

OY 414 NLVSGSKPSEATCKNVFNNAEHVTDPCT 443
DB 407 KIRG-----GPAACKMANVNGQIVSPKCS 431

RESULT 10

AAAG41323
ID AAAG41323 standard; protein; 463 AA.

XX AAAG41323;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 51399.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0125264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134768P.

PR 20-MAY-1999; 99US-0135353P.

PR 21-MAY-1999; 99US-0135529P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137528P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 14-JUN-1999; 99US-0138847P.

PR 16-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 21-JUN-1999; 99US-0139763P.

PR 22-JUN-1999; 99US-0139817P.

PR 23-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 24-JUN-1999; 99US-0140354P.

PR 28-JUN-1999; 99US-0140695P.

PR 29-JUN-1999; 99US-0140823P.

PR 30-JUN-1999; 99US-0140991P.

PR 01-JUL-1999; 99US-0141287P.

PR 02-JUL-1999; 99US-0141842P.

PR 02-JUL-1999; 99US-0142154P.

PR 06-JUL-1999; 99US-0142055P.

PR 08-JUL-1999; 99US-0142390P.

PR 09-JUL-1999; 99US-0142803P.

PR 12-JUL-1999; 99US-0142920P.

PR 13-JUL-1999; 99US-0143277P.

PR 15-JUL-1999; 99US-0143624P.

PR 16-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 19-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 21-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 22-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 23-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 26-JUL-1999; 99US-0145224P.

PR 27-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145919P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 05-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 06-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 09-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

PR 10-AUG-1999; 99US-0147935P.

PR 11-AUG-1999; 99US-0148171P.

PR 12-AUG-1999; 99US-0148319P.

PR 13-AUG-1999; 99US-0148565P.

PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	13-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154033P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155133P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157533P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	23-OCT-1999;	99US-0160989P.
PR	23-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

	Query Match	47.2%	Score 1127;	DB 3;	Length 463;	
	Best Local Similarity	49.6%;	Pred. No. 1.5e-86;			
	Matches 223;	Conservative 71;	Mismatches 116;	Indels 40;	Gaps 6;	
Qy	7 SILLIIIPASSISCTGNSVID-----DVLFPQVYNNIIEGEAFHDFQAYLXL 55					
	:: :::: :::::					
Dd	41 AVFLCVLLIMSLCKRLSSNVDDGYCHEGSGPESDLKLKANDDVL-----SLISSD 91					
Qy	56 SKNISSNNNIDKVDKNIGIKIVNLVSFGAKGDOKTYDNAIFEQANMEACHSRTPVOFVPVK 115					
	::::: ::::: ::::: :					

```

Db      92  ETTLEBS-----TVSVPFGAKGDKGTDDTQAFKAWKKACSTGVTTFVLVPK 139
QY      116  NKNYLLKOTFFSGPCRRSISVKI PSLSEASSKISYDKORLMIAPDSYQNLVYGG--GT 173
Db      140  GKTLLKSTRFRPQCSKLTNPFQILGTLLGASTRSPDYKKNHMLILEDNVNLISDIGSGTGI 199
QY      174  INNGQVWMPSSCKINKSLPCRDATLALTFWNCCKLKNYLKSNRAOQIHKFPCSTNVV 233
Db      200  INNGKTMWONSCKIDKSRCTKAPALTALYNLKNLVKNLRVKNRAOQIOLSIKKCNVE 255
QY      234  ASNLIMNASKSPNTDGVHVSNTQYIOISDITIGTGDDCISIVSAGSQVQATNTCGPGH 293
Db      260  VSNVEITAPEDSPNTDGIHTTQNIKRNNSDSIGDDCISIEBGTQNLQFDLTGCPGH 319
QY      294  GISIGSGSGNSEAVYSNTTVNEAKIIGAENGVRIKTVOGSGGASNIKPLNVMQVY 353
Db      320  GISIGSLGDDNSKAYVSGINDGAKFPESDNRVRIKTYQGGSGTAHAKNKFQINMEHVN 379
QY      354  PIITDQYCDRVEPCTIQPFAVQKRVVYENIKGTSATKVAIKPDCSTNFPCEGIENI 413
Db      380  PIITDQYCDK-DKCEDQESAQVQKRVVYKNISGTSATDVAITLNCSEKYPQCGIVLENV 438
QY      414  NLVGSQKBEALCTCKVHFNMAEHTPHCT 443
Db      439  KING-----GTASCKNANVKNQTVSPKCS 463

```

RESULT 11	
AAAG41325	
ID	AAAG41325 standard; protein; 415 AA.
XX	
AC	AAAG41325;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 51401.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.

PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134766P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0136622P.
PR	25-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136021P.
PR	01-JUN-1999;	99US-0136782P.
PR	03-JUN-1999;	99US-0137222P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	14-JUN-1999;	99US-0138847P.
PR	16-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	22-JUN-1999;	99US-0139817P.
PR	23-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140554P.
PR	28-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	30-JUN-1999;	99US-0140991P.
PR	01-JUL-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	02-JUL-1999;	99US-0142154P.
PR	06-JUL-1999;	99US-0148055P.
PR	08-JUL-1999;	99US-0142390P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143624P.
PR	14-JUL-1999;	99US-0143542P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144684P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145081P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145148P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	10-AUG-1999;	99US-0147935P.
PR	11-AUG-1999;	99US-0148171P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151080P.
PR	31-AUG-1999;	99US-0151303P.
PR	01-SEP-1999;	99US-0151438P.
PR	07-SEP-1999;	99US-0151930P.
PR	10-SEP-1999;	99US-0152363

OS Brassica napus; cv. Rafal.
 XX W09630529-A1.
 XX 03-OCT-1996.
 PD 29-MAR-1996; 96WO-GB000757.
 XX 31-MAR-1995; 95GB-0006684.
 XX (NICK-) NICKERSON BIOCHEM LTD.
 PA Roberts JA, Coupe SA, Jenkins ES;
 XX WPI: 1996-455374/45.
 DR N-PSDB; AATJ33994.
 XX
 PT Control of seed pod dehiscence - using polygalacturonase or nucleic acid
 XX sequences derived from polygalacturonase gene.
 XX
 PS Example 1; Fig 1; 36pp; English.
 CC A polygalacturonase (PG) (AAW04248) of oilseed rape cv. Rafal seed pods
 CC is useful for controlling dehiscence. Manipulation of the enzyme's
 CC activity can influence the timing of dehiscence. A cDNA clone (AATJ3994)
 CC coding for the PG was isolated from a cDNA library of the rape seed pod
 CC dehiscence zone. PG nucleic acids (patric. antisense) can be used to
 CC regulate dehiscence in crop plants. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX
 XX Sequence 433 AA:

Query Match 46.2%; Score 1103; DB 2; Length 433;
 Best Local Similarity 49.8%; Pred. No. 1.4e-84;
 Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;
 QY 7 SILLIIIFASSISTERSNVID-----DNLFKQVYDNIIEQEFADHPOAVLYSTYKNIYES 61
 Db 9 AITLCVLLMLACCOALSSNVDDGYGHEGDSFET--DSLILKLNDDVLTLSKSDRPTTES 66
 QY 62 NNNIDKVDKNGIKVINVLSPFGAKGDKTYDNIAPFOAMNEACSRTPVQFVPPKNNYYLL 121
 Db 67 S-----TVSVNPFAGKDGKTDPTQAFKAKMKKACSTNGVTTLIRPKGYLL 114
 QY 122 KOITSGPCRSISYKIFGSLPSSASKISDY-KDRRLMIAPDSYQNTLVGSG--GTINGNG 178
 Db 115 KSIRRGCKSLRFPQILGTLSASTKSDYNDKNHMLLEDVNNLSIDGSGAGIVDNG 174
 QY 179 QVWPPSSCKINKSLPCRDAPALTLPNNCKYLKYNMLKSKNAQOIHIKFECSCTNVVANSML 238
 Db 175 KIMMONSCKIDKSPCTKAPTALTLYNLNLANVKNLRVNAQOIQISIEKNSVDVKNVK 234
 QY 239 INASAKSPNDGVAHVSTQYIQTSDTITIGTDDCISIVGSONVOATNITGPGHGISTG 298
 Db 235 ITAGDSPNDGHIHVAIKNIRISNSDIGTDCISIEDGSONVQINDLTGPGHGISTG 294
 QY 299 SLGSGNSEAVSVNTVNEAKITGAENGVRITKMOGSGQASNIKFLANVEMQDVXPIIID 358
 Db 295 SLGDNSKAVSGINVDATLSETDNGVRITKYOGSGTAKNIRKQNTRMNVNKPPIIID 354
 QY 359 QNYCDREPCIOQFSAVQVKNVYENIKGTSATKYAIFDCSTNFPCEGIIMENINLVGE 418
 Db 355 QNYCDR-DKCEQOESAVQVNNVYRNIGTSTADVAIMFNCSVYKPCOGIYLVENVIK- 412
 QY 419 SGKPSATCKVNFNNAEHVTPHC 442
 Db 413 ----GKASCKVNVKDKGTVSPKC 432

RESULT 14
 AAAY42649
 ID AAAY42649 standard; protein; 433 AA.

AC AAAY42649;
 XX 10-JAN-2000 (first entry)
 DT Brassica napus Sac66 protein putative sequence.
 DE Signal transduction protein; dehiscence; male sterile plant; DZ2 gene;
 XX shatter resistance; oilseed rape; Sac66 protein.
 KW Brassica napus.
 OS W09949046-A1.
 XX 30-SEP-1999.
 PD 22-MAR-1999; 99WO-GB000905.
 XX 20-MAR-1998; 98GB-0006113.
 XX (BIOG-) BIOGENMA UK LTD.
 PA Wyatt P, Roberts JA, Whiteclaw C;
 XX WPI: 1999-580449/49.
 DR N-PSDB; AA222980.
 XX
 PT A nucleic acid encoding a signal transduction protein involved in plant
 XX dehiscence, useful for producing shatter resistant male sterile plants.
 XX
 PS Example 6; Fig 15; 71pp; English.
 CC The invention provides a nucleic acid encoding a signal transduction
 CC protein involved in the process of dehiscence. The nucleic acids and
 CC proteins are useful for regulating or controlling dehiscence of a pod or
 CC an anther in a plant, useful in the production of male sterile plants.
 CC The methods, etc. may be used in production of shatter resistance or
 CC shatter-delayed plants such as oilseed rape (Brassica napus). The present
 CC sequence represents a B. napus Sac66 protein putative sequence
 XX
 XX Sequence 433 AA:

Query Match 46.2%; Score 1103; DB 2; Length 433;
 Best Local Similarity 49.8%; Pred. No. 1.4e-84;
 Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;
 QY 7 SILLIIIFASSISTERSNVID-----DNLFKQVYDNIIEQEFADHPOAVLYSTYKNIYES 61
 Db 9 AITLCVLLMLACCOALSSNVDDGYGHEGDSFET--DSLILKLNDDVLTLSKSDRPTTES 66
 QY 62 NNNIDKVDKNGIKVINVLSPFGAKGDKTYDNIAPFOAMNEACSRTPVQFVPPKNNYYLL 121
 Db 67 S-----TVSVNPFAGKDGKTDPTQAFKAKMKKACSTNGVTTLIRPKGYLL 114
 QY 122 KOITSGPCRSISYKIFGSLPSSASKISDY-KDRRLMIAPDSYQNTLVGSG--GTINGNG 178
 Db 115 KSIRRGCKSLRFPQILGTLSASTKSDYNDKNHMLLEDVNNLSIDGSGAGIVDNG 174
 QY 179 QVWPPSSCKINKSLPCRDAPALTLPNNCKYLKYNMLKSKNAQOIHIKFECSCTNVVANSML 238
 Db 175 KIMMONSCKIDKSPCTKAPTALTLYNLNLANVKNLRVNAQOIQISIEKNSVDVKNVK 234
 QY 239 INASAKSPNDGVAHVSTQYIQTSDTITIGTDDCISIVGSONVOATNITGPGHGISTG 298
 Db 235 ITAGDSPNDGHIHVAIKNIRISNSDIGTDCISIEDGSONVQINDLTGPGHGISTG 294
 QY 299 SLGSGNSEAVSVNTVNEAKITGAENGVRITKMOGSGQASNIKFLANVEMQDVXPIIID 358
 Db 295 SLGDNSKAVSGINVDATLSETDNGVRITKYOGSGTAKNIRKQNTRMNVNKPPIIID 354
 QY 359 QNYCDREPCIOQFSAVQVKNVYENIKGTSATKYAIFDCSTNFPCEGIIMENINLVGE 418
 Db 355 QNYCDR-DKCEQOESAVQVNNVYRNIGTSTADVAIMFNCSVYKPCOGIYLVENVIK- 412

This page Blank (uspto)

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 06:59:05 ; Search time 41 Seconds
(without alignments)
1072.465 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MWIQRNILLILLIIFASIS.....VTPHCTSLSEDEALYNY 457

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	457	2	A25534 polygalacturonase
2	1142	47.8	462	1	S31195 polygalacturonase
3	1128	47.2	461	2	T08215 polygalacturonase
4	1127	47.2	431	2	T46187 polygalacturonase
5	1115.5	46.7	426	2	H84846 probable polygalac
6	1108	46.4	458	2	S71523 polygalacturonase
7	1091	45.6	460	2	T17011 polygalacturonase
8	898.5	37.6	459	2	D96833 hypothetical prote
9	840.5	35.2	468	2	H86728 probable polygalac
10	796.5	33.3	514	2	JC7100 polygalacturonase
11	787.5	32.9	434	2	A96609 probable polygalac
12	787.5	32.9	514	2	S48720 Cys 1 II protein -
13	787	32.9	514	2	JC2498 second major aller
14	764	32.0	507	2	JC7366 un a 2 protein -
15	749.5	31.4	1161	2	B86368 protein F28C11.9 l
16	741.5	31.0	491	2	B86155 probable polygalac
17	736	30.8	423	2	T08213 polygalacturonase
18	728	30.5	405	2	D84871 probable polygalac
19	727	30.4	392	2	S57806 polygalacturonase
20	718.5	30.1	395	2	T05906 probable polygalac
21	718	30.0	435	2	T48618 probable polygalac
22	715	29.9	392	2	T00669 probable polygalac
23	711.5	29.8	387	2	T04330 polygalacturonase
24	710.5	29.7	365	2	T04319 polygalacturonase
25	708.5	29.6	422	2	S34199 polygalacturonase
26	704.5	29.5	387	2	T04352 polygalacturonase
27	704	29.5	374	2	T04672 probable polygalac
28	704	29.5	374	2	B85421 hypothetical prote
29	702.5	29.4	394	2	G86190

30	701.5	29.4	542	2	H86239 protein F20B24.8 l
31	698.5	29.2	452	2	C85024 probable polygalac
32	696	29.1	384	2	E84871 probable polygalac
33	690	28.9	540	2	B96631 probable polygalac
34	688.5	28.8	492	2	C96521 protein F21D18.18
35	687.5	28.8	393	2	S40123 polygalacturonase
36	687	28.7	383	2	T47809 polygalacturonase
37	684	28.6	397	2	F96680 F514.10 (imported
38	683	28.6	394	2	F86190 hypothetical prote
39	680.5	28.5	394	2	T00668 probable polygalac
40	678.5	28.4	445	2	S34266 polygalacturonase
41	673.5	28.2	387	2	T07591 polygalacturonase
42	673	28.2	444	2	S34200 polygalacturonase
43	667	27.9	407	2	S52006 polygalacturonase
44	662.5	27.7	664	2	A84742 probable polygalac
45	659	27.6	414	2	E85204 polygalacturonase-

ALIGNMENTS

RESULT 1

A25534 polygalacturonase (EC 3.2.1.15) precursor - tomato
N:Alternate names: pectinase; poly[1,4-alpha-D-galacturonide]glucanohydrolase
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: A25534; J0156; S06340; S02101; S08504
R:Grierson, D.; Tucker, G.A.; Keen, J.; Ray, J.; Bird, C.R.; Schuch, W.
Nucleic Acids Res. 14, 8595-8603, 1986
A:Title: Sequencing and identification of a cDNA clone for tomato polygalacturonase.
A:Reference number: A25534; PMID:87066731; PMID:3786135
A:Accession: A25534
A:Molecule type: mRNA
A:Residues: 1-457 <GR1>
A:Cross-references: UNIPROT:P05117; GB:X04583; NID:g19291; PIDN:CAA28254.1; PID:g19292
R:Delipenna, D.; Bennett, A.B.
Plant Physiol. 86, 1057-1063, 1988
A:Title: In vitro synthesis and processing of tomato fruit polygalacturonase.
A:Reference number: J0156
A:Accession: J0156
A:Molecule type: mRNA
A:Residues: 1-115
A:Cross-references: GB:M20269; NID:g170470; PIDN:AAA34177.1; PID:g170471
A:Experimental source: fruit
R:Sheehy, R.E.; Pearson, J.; Brady, C.J.; Hlatk, W.R.
Mol. Gen. Genet. 208, 30-36, 1987
A:Title: Molecular characterization of tomato fruit polygalacturonase.
A:Reference number: S06340
A:Accession: S06340
A:Molecule type: mRNA
A:Residues: 1-457 <SHE>
A:Cross-references: EMBL:X05656; NID:g19297; PIDN:CAA29148.1; PID:g19298
A:Note: part of this sequence, including the amino end of the mature protein, was confirm
R:Rose, R.E.; Houck, C.M.; Monson, E.K.; DeDeus, C.E.; Sheehy, R.E.; Hlatk, W.R.
Nucleic Acids Res. 16, 7191, 1988
A:Title: The nucleotide sequence of the 5' flanking region of a tomato polygalacturonase
A:Reference number: S02101; PMID:88303350; PMID:3405763
A:Accession: S02101
A:Molecule type: DNA
A:Residues: 1-93 <ROS>
A:Cross-references: EMBL:X07410; NID:g19295; PIDN:CAA30308.1; PID:g19296
R:Bird, C.R.; Smith, C.J.S.; Ray, J.A.; Moureau, P.; Bevan, M.W.; Bird, A.S.; Hughes, S.;
Plant Mol. Biol. 11, 651-662, 1988
A:Title: The tomato polygalacturonase gene and ripening-specific expression in transgenic
A:Reference number: S08504
A:Accession: S08504
A:Molecule type: DNA
A:Residues: 1-457 <BIR>
A:Cross-references: EMBL:X14074; NID:g19305; PIDN:CAA32235.1; PID:g295813
C:Comment: The polygalacturonase, a single cell wall enzyme, is the major enzyme mediactr
C:Superfamily: polygalacturonase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-71/Domain: propeptide #status predicted <PRO>
F:72-444/Product: polygalacturonase 2A #status experimental <MAT>

Query Match

100.0%; Score 2390; DB 2; Length 457;
Best Local Similarity 100.0%; Pred. No. 5, 6e-163;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQNSILLIITIFASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVLYSLSKNIE 60
DB 1 MVIQNSILLIITIFASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVLYSLSKNIE 60
QY 61 SNNIDVADKNGIKVINLVSPGAKGDKTYDNIAPQAMNEACSSRTFQVFPKKNYL 120
DB 61 SNNIDVADKNGIKVINLVSPGAKGDKTYDNIAPQAMNEACSSRTFQVFPKKNYL 120
QY 121 LKQITFSGCRSSISVKIFGSLBASISKIDYKDRRLWIAPDSYONLVGGGGTINGNGQV 180
DB 121 LKQITFSGCRSSISVKIFGSLBASISKIDYKDRRLWIAPDSYONLVGGGGTINGNGQV 180
QY 181 WMPSSCKINKSLPCRDAPALTALTFWNCNKLKNNLKSNAQOIHIFESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTALTFWNCNKLKNNLKSNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISPTIITGTGDDCISIVSGSONVOATNITCGGHSIGSL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISPTIITGTGDDCISIVSGSONVOATNITCGGHSIGSL 300
QY 301 GSGNSEAVSNVNVNEAKITGAENGVRITWOGGSGQASNKFLVEMQDVKYPPIIIDON 360
DB 301 GSGNSEAVSNVNVNEAKITGAENGVRITWOGGSGQASNKFLVEMQDVKYPPIIIDON 360
QY 361 YCRVBPCTIQPSAVQVKNVYENIKGSATKVAIKFDCSTNPFCEGIIMENINIVGEG 420
DB 361 YCRVBPCTIQPSAVQVKNVYENIKGSATKVAIKFDCSTNPFCEGIIMENINIVGEG 420
QY 421 KSEATCKKVNHNMAEHTPHCTSLTSEDEALLVNY 457
DB 421 KSEATCKKVNHNMAEHTPHCTSLTSEDEALLVNY 457

```

RESULT 2

S31195
polygalacturonase (EC 3.2.1.15) - avocado
C:Species: Persea americana (avocado)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
R:Accession: S31195; S28072
P:Dojico, B.; Lowe, A.L.; Wilson, I.D.; Merodio, C.; Grierson, D.
A:Title: Cloning and characterization of avocado fruit mRNAs and their expression during
A:Reference number: S31195; MUID:93184201; PMID:8095163
A:Molecule type: mRNA
A:Residues: 1-462 <DOP>
A:Cross-references: UNIPROT:Q02096; EMBL:X66426
R:Dojico, B.
submitted to the EMBL Data Library, May 1992
A:Reference number: S28072
A:Accession: S28072
A:Molecule type: mRNA
A:Residues: 1-181, 'S', 183-462 <DOP>
A:Cross-references: EMBL:X66426; NID:922630; PID:CAA7055.1; PID:922631
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match

Best Local Similarity 47.8%; Score 1142; DB 1; Length 462;
Matches 224; Conservative 67; Mismatches 106; Indels 8; Gaps 5;

```

QY 28 DDLNFQVYDNLLEQFAHDFQAVLYSLSKNIESNNIDVADKNGIKVINLVSPGAKG 87
DB 40 DDLNFQVYDNLLEQFAHDFQAVLYSLSKNIESNNIDVADKNGIKVINLVSPGAKG 87

```

```

QY 88 KTYNIAPEQAMNEACSSRTFQVFPKKNYLKQITFSGCRSSISVKIFGSLBASISK 147
DB 97 -TDTKAFERKAMDACSSSVL--VPEKNKYLKQITFSGCRSPDRVIRITIASSD 153
QY 148 ISDY--KDRRLWIAPDSYONLVGGGGTINGNGQVWMPSSCKINKSLPCRDAPALTAL 205
DB 154 OSDVGHNKNRIEFDISNLTLEGGGTINGNGTWDSSCKRKLPCSKAPALTALFES 213
QY 206 CKRLKNNLKSNAQOIHIFESCTNVVASNLMINASASPNTDGVHVSNTQYIQISPTI 265
DB 214 CKRLKNNLKSNAQOIHIFESCTNVVASNLMINASASPNTDGVHVSNTQYIQISPTI 273
QY 266 IGTGDDCISIVSGSONVOATNITCGGHSIGSLGSGNSEAVSNVNVNEAKITGAENG 325
DB 274 IGTGDDCISIVSGSONVOATNITCGGHSIGSLGSGNSEAVSNVNVNEAKITGAENG 333
QY 326 VRIKTWOGGSGQASNKFLVEMQDVKYPPIIIDONYCRVBPCTIQPSAVQVKNVYENI 385
DB 334 LRITWOGGSGQASNKFLVEMQDVKYPPIIIDONYCRVBPCTIQPSAVQVKNVYENI 393
QY 386 KGSATVVAIKFDCSTNPFCEGIIMENINIVGEGSEATCKVNY 430
DB 394 RGTSAEVAVKFDCSKSPQCYIVGNINLVNGKETTWSGSI 438

```

RESULT 3

T08215
polygalacturonase (EC 3.2.1.15) 3 precursor - muskmelon
C:Species: Cucumis melo (muskmelon)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
R:Hadfield, K.A.; Rose, J.K.; Yaver, D.S.; Berka, R.M.; Bennett, A.B.
Plant Physiol. 117, 363-373, 1998
A:Title: Polygalacturonase gene expression in ripe melon
A:Reference number: 216403; MUID:98289082; PMID:9625689
A:Accession: T08215
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-461 <HAD>
A:Cross-references: UNIPROT:Q81246; EMBL:AF062467; NID:93320461; PID:93320462
C:Genetics:
A:Gene: MRC3
C:Superfamily: polygalacturonase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-461/Product: polygalacturonase #status predicted <MAT>

Query Match

Best Local Similarity 47.2%; Score 1128; DB 2; Length 461;
Matches 224; Conservative 79; Mismatches 129; Indels 14; Gaps 7;

```

QY 7 SILLIITIFASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVLYSLSKNIESNNI 65
DB 17 TFLFVNVNFDRTSTCFSGYPDVNPPLSISGKEDVQGHYHSSSEFSMLRT--RL 74
QY 66 DKVDKNGI--KVINLVSPGAKGDKTYDNIAPQAMNEACSSRTFQVFPKKNYLK 122
DB 75 EKVVSSPLASPEIRFVNDYDYGAGDGED--DEAFRETKDKACSTNAT--PLVPCDRYHLK 132
QY 123 QITFSGCRSSISVKIFGSLBASISKIDY--KDRRLWIAPDSYONLVGGGGTINGNGQV 181
DB 133 PTFSGPCNSPLTFQIEGTVAAPHISIDYKDRRLWIAPDSYONLVGGGGTINGNGQV 192
QY 182 WPSCKINKSLPCRDAPALTALTFWNCNKLKNNLKSNAQOIHIFESCTNVVASNLMIN 241
DB 193 WINSCKVKKTLPCREATAVTFQCTNLREVEGLRFRNAQOHLSTGRCKNNKALNLTIVA 252
QY 242 SAKSPNTDGVHVSNTQYIQISPTIITGTGDDCISIVSGSONVOATNITCGGHSIGSL 301
DB 253 PGNBPNTDGHVHTGTFIVIKNCLIMTGDGDCISIVSGSKNVRAGITCGGHSIGSL 312
QY 302 GGNSEAVSNVNVNEAKITGAENGVRITWOGGSGQASNKFLVEMQDVKYPPIIIDON 361

```

Db 313 AGSEAEVSNVVDYTAKEFGSTNGVRIRKTVQGGKGYAQNIIIFONIVMDVNTNPIIINONY 372
Qy 362 CDRVEPCIQGFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLV----- 416
Db 373 CDQKEPCTQQADVAVSNVMTQNIKRTSASEVAIKFDCSVPCQGLDIDINLVHKGKN 432
Qy 417 GEGSKPEATCTKVVHFNNAEHVTPHC 442
Db 433 DDKSAQAEASCKVKNKRGKRVSPQC 458

RESULT 4

T46187
polygalacturonase (EC 3.2.1.15) precursor [similarity] - Arabidopsis thaliana
N/alternate names: protein T8H10.110
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46187, T50674
R/Bones, V.; Rechmann, S.; Borkova, D.; Ansgore, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.H.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23014
A/Accession: T46187
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-431 <BEN>
A/Status: preliminary
A/Accession: T50674
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Residues: 1-431 <JEN>
A/Cross-references: EMBL:AF037367; PIDN:AAC98923.1
A/Experimental source: cultivar Landsberg erecta
C/Genetic:
A/Map position: 3
A/Intons: 85/3; 129/3; 186/3; 193/3; 263/1; 290/2; 326/3; 364/3
C/Superfamily: polygalacturonase
C/Keywords: glycosidase; hydrolase
F/1-27/Domain: signal sequence #status predicted <SIG>
F/28-431/Product: polygalacturonase #status predicted <MAT>

Query Match

Best local similarity 49.6%; Score 1127; DB 2; Length 431;
Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

Qy 7 SILLIIIFASSISTCRSNVID-----DNLFKQYVDNITLBEFAHDFQAYLSYL 55
Db 9 AVLCLVLMISLCKALSSNVDDGHEHDSFESDSLKLNNDVL-----SLISSD 59
Qy 56 SKRIEENNNDKYDKNGIKVINLVSPGAKGDKTYDNIAEFOAMNEACSRTPVOFVVPK 115
Db 60 ETTLEAS-----TVSVNFGAKGDKTDDTQAFKAKMKKACSTNGVTFVVPK 107
Qy 116 NKVYLLKQITFGSPCRSSISVKIFGSLSEASSKISDYDRRLMIAFDSVNLVVGCG--GT 173
Db 108 GKTYLLKSTRFRGPKSLRNFQILGTLASASTKSDYDKKHMWLLLEEDVNNLSIDGGSTGI 167
Qy 174 INNGGVWVPSCKINKSLCRDAPTLTFMNCCKNLKVNLSKSNAAQIHIKESCTNVV 233
Db 168 INNGKTMWNSCKIDSKPCTKAPVALLTYLNELNVLNVRKNAQOIQISIECKNVE 227
Qy 234 ASNLMINASAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSNVQATNITCGPGH 293
Db 228 VSNVEIRIAPDPSPTDGIHTNTQNIKRVNSDICTGDCISIEBGTQNLIDPDLTGCPGH 287
Qy 294 GISIGSGNSSEAVYNSVTNVEAKITGAENGVRIRKTVQGGSGOASNIKFLVEMQDVY 353
Db 288 GISIGSGDNDSKAVYSGIINVDGAKFSESDNGVRIKTYQGGSGTAKNIKFNIRMEVKN 347

Qy 354 PIIDQNYCDRVEPCIQGFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENI 413
Db 348 PIIDDDYDCK-DKCDQSSAVQVKNVYENIKGTSATDVAIILNCSBKPPCGIYLENV 406
Qy 414 NLVBSGKPEATCTKVVHFNNAEHVTPHCT 443
Db 407 KIKG-----GTASCKNANVNGCTVSPKCS 431

RESULT 5

H84846
probable polygalacturonase [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: H84846
R/Fin, X.; Kaul, S.; Rounale, S.D.; Shea, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; NCID:20083487; PMID:10617197
A/Accession: H84846
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-426 <STO>
A/Cross-references: UNIPROT:O22935; GB:AE002093; NID:G2335094; PIDN:AAC02763.1; GSPDB:GNC
C/Superfamily: polygalacturonase

Query Match

Best local similarity 49.8%; Score 1115.5; DB 2; Length 426;
Matches 222; Conservative 74; Mismatches 117; Indels 33; Gaps 7;

Qy 6 NSILLIIIFASSISTCRSNVIDNLFKQY-----DNLEDFAHDFQAYLSYLSKITE 60
Db 6 NLTVFLMALMLFSGCKASRISPNVYDHSYKFKSDSLIKR--EDITGRSFFVASLR 63
Qy 61 SNNNIRKVDKNGIKVINLVSPGAKGDKTYDNIAEFOAMNEACSRTPVOFVVPKKNYL 120
Db 64 TPT-----TVSVDPFGAKGDKTDDTQAFVNAWKACSSNGAVNLVLPKNTYL 112
Qy 121 LKQITFGSPCRSSISVKIFGSLSEASSKISDYDRRLMIAFDSVNLVVGCG--GTINNG 178
Db 113 LKSIQTLGPNSLTVQIFGTLASQKRSYKDIKIMFEDGNNLSVDSGDTGVVDGNG 172
Qy 179 QVWVPSCKINKSLPCRDAPTLTFMNCCKNLKVNLSKSNAAQIHIKESCTNVVASNLM 238
Db 173 ETWQNSCKRNK-----KALTFFNSKSLIVKMLKVRNAQOIQISIEKSNVQSNV 225
Qy 239 INASAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSNVQATNITCGPGHGISIG 298
Db 226 VTPAPDSPTDGIHTNTQNIKRVNSDICTGDCISIESSQVQVNDITGCPGHGISIG 285
Qy 299 SLGSGNSEAVYNSVTNVEAKITGAENGVRIRKTVQGGSGOASNIKFLVEMQDVYPIID 358
Db 286 SLGDDNSKAFVSVTVYDGAALSTGDNGVRIRKTYQGGSGTRSNIIIFONIVMDVNTNPIIID 345
Qy 359 QNYCDRVEPCIQGFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGE 418
Db 346 QDYCDK-SKCTTEKSAVQVKNVYRISGTSASENNITFCGKNYCCGIVLDRVNIKG-- 403
Qy 419 GSKPSEATCTKVVHFNNAEHVTPHCTS 444
Db 404 ---GRATCTNANVVDGKAVLPQCN 425

RESULT 6

S71523
polygalacturonase (EC 3.2.1.15) [similarity] - peach
N/alternate names: endopolygalacturonase

C:Species: Prunus persica (peach)
 C:Date: 19-Mar-1997 #sequence revision 19-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S71523; S71524; S71525
 R:Lee, B.; Speirs, J.; Gray, J.; Brady, C.J.
 A:Submitted to the EMBL Data Library, April 1994
 A:Description: Homologues to the tomato endopolygalacturonase gene in the peach genome.
 A:Reference number: S71523
 A:Accession: S71523
 A:Molecule type: DNA
 A:Residues: 1-458 <LEB>
 A:Cross-references: UNIPROT:Q43063; EMBL:X77231; NID:G479087; PIDN:CAA54448.1; PID:G4790
 A:Experimental source: cultivar Maravilla
 A:Note: this is a revision to the sequence from reference S71524
 R:Lee, B.; Speirs, J.; Gray, J.; Brady, C.J.
 A:Submitted to the EMBL Data Library, January 1994
 A:Description: Homologues to the tomato endopolygalacturonase gene in the peach genome.
 A:Reference number: S71524
 A:Accession: S71524
 A:Molecule type: DNA
 A:Residues: 53-92, 'I', '94-149, 'T', '151-185, 'W', '187-196, 'THA', '200, 'ESLS', '206-207, 'TNT', '212-
 A:Cross-references: EMBL:X77231
 A:Experimental source: cultivar Maravilla
 A:Note: this sequence has been revised in reference S71523
 R:Lee, B.; Speirs, J.; Gray, J.; Brady, C.J.
 A:Plant Cell Env. 13, 513-521, 1990
 A:Title: Homologues to the tomato endopolygalacturonase gene in the peach genome.
 A:Reference number: S71525
 A:Accession: S71525
 A:Molecule type: DNA
 A:Residues: 265-458 <LEW>
 A:Cross-references: EMBL:X77231
 A:Experimental source: cultivar Maravilla
 C:Genetics:
 A:Insertions: 107/3; 150/3; 204/3; 213/3; 283/1; 310/2; 386/3
 C:Superfamily: polygalacturonase
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 46.4%; Score 1108; DB 2; Length 458;
 Best Local Similarity 48.9%; Pred. No. 1.8e-71;
 Matches 226; Conservative 82; Mismatches 126; Indels 28; Gaps 8;

```

QY 1 MWIORSILILIIIFASSISITCRSNVIDNLFKQVYDNIIEQFPAHDFQAYLSYLSK----56
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MLKQLMLSPFVVVVFSTTSCGSPFOVNLHSHVVDKESGYNSRAYPSY-ID 59
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 57 -----KNIESNNIDKY---DKNGIKVINVLSPGAKGDKGTYNINAF 103
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 59 VKEMEFIRPRAQLFESSRLERAGSKSSSVKTIIVANFGAKNGAD-DTAEFKAWKAC 117
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 104 SSRTPVGVPRKNKRYLLKQITFSGPCSSISVKIFGSLASAKSIDYKDRRLMIAFDSV 163
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 118 SSNGAIVLVPPQ-KTYLVNPLIEFSGPCSKSHLTMOIYGTIEASDSDRSVYKDVTHMLFDNV 176
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 164 QNLVVGSGGTINGNGQVWPPSSCKINKSLPCRD-APFALTFPNNCKRLKNNLSKKAQOI 222
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 QSLVVGSGGTINGNGNEMWENSCKRKPQPCNFOAFLATVFNKCNLVKNIKIDDAQOM 236
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 223 HIKFESCTNVVASSLMNINAKSPNTDGVAVSTQYIQISDTIIGTGDDCISIVSGSNV 282
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 HVKRNCKNVEASHLTATAPEDSPNTDGIHNTKNIITISSSVITGTDDCISIVSGSRV 296
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 283 QATNITCGPHGISIGSLGSGNSEAVYNSVTVEAKTIGANGVRIKTKWGGSGGASNIK 342
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 297 QATDITCGPHGISIGSLGSDNANDHVGVPFNGAKISTSGVGIKTKWGGSGGASNTV 356
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 343 FLVNEQMDVKKPIIITDNYCD-RVEPCIQGFAVQVKNVYENIKGTSATKVAIKFDDST 401
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 357 FQVNEQMDVKNPIIITDNYCDKHNKDCRQSAVQVKNVLYQNIKRTGTSASDALTFFNCQ 416
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 402 NFPCESGIMENINLVGSGSPSEATCKNVHFNNAHVTPHC 443
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 417 SVFCQGVIVQNIQL-----QNRACNNVKAIVKAVSPRCS 453
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 7
 T17011
 polygalacturonase (EC 3.2.1.15) - apple tree
 N:Alternate names: poly (1,4-a-D-galacturonide) glycan hydrolase
 C:Species: Malus domestica (apple tree)
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 R:Atkinson, R.G.
 A:Plant Physiol. 105, 1437-1438, 1994
 A:Title: A cDNA clone for endopolygalacturonase from apple.
 A:Reference number: Z18649; MUID:95062722; PMID:7972500
 A:Accession: T17011
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-460 <ATK>
 A:Cross-references: UNIPROT:P48978; EMBL:L27743; NID:G456091; PIDN:AAA74452.1; PID:G4560
 C:Function:
 A:Description: catalyzes hydrolysis of 1,4-alpha-D-galactosiduronic linkages in galactur
 A:Pathway: polysaccharide degradation
 C:Superfamily: polygalacturonase
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 45.6%; Score 1091; DB 2; Length 460;
 Best Local Similarity 49.3%; Pred. No. 2.9e-70;
 Matches 231; Conservative 76; Mismatches 126; Indels 36; Gaps 12;

```

QY 1 MWIORSILILIIIFASSIST--CRSNVIDNLFKQVYDNIIEQFPAHDFQAYLSYLSK 57
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MLKQLMLSPFVVVVFSTTSCGSPFOVNLHSHVVDKESGYNSRAYPSY-ID 59
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 58 NIE-----SNNNIDKVDKNGI-----KVYNLSPGAKGDKGTYNINAF 95
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 60 TIEGKVMELIRPRTQLFSSRLKNTI-TGGIATSSAPAKTISVDDFGAKNGAD-DTQAF 117
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 96 EQANWEACSSRTPQVFPVPRKNKRYLLKQITFSGPCSSISVKIFGSLASAKSIDYKDR 155
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 118 VKAMWAAKSSSGANVAVLPQ-KTYLVNPLIEFSGPCSKSHLTMOIYGTIEASDSDRSVYKDV 176
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 156 LMIAFDSVQNLVVGSGGTINGNGQVWPPSSCKINKSLPCRD-APFALTFPNNCKRLKNNLS 214
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 HMLIFDVQNLVVGSGGTINGNGNIMWKSCKIKPQPCGTAFLATVFNKCNLVKNI 236
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 215 KSKAAQOIHIFKPSCTNVVASSLMNINAKSPNTDGVAVSTQYIQISDTIIGTGDDCIS 274
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 NIDPAQOIHIVFQNCINVAQSLTATAPEDSPNTDGIHNTKNIITISSSVITGTDDCIS 296
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 275 IVSGSNVQATNITCGPHGISIGSLGSGNSEAVYNSVTVEAKTIGANGVRIKTKWGG 334
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 297 IVSGSGRVQATDITTCGPHGISIGSLGSDGSDHDSGVFNGAKISTSGVGIKTKWGG 356
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 335 SGOASNIKFNEMQVKKPIIITDNYCD-RVEPCIQGFAVQVKNVYENIKGTSATKV 393
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 357 SGAATNIVQNVQMDVKNPIIITDNYCDHNTKCKCKQKSAVQVKNVLYQNIKRTGTSASD 416
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 394 AIFDCSTNFPCEGIMENINLVGSGSPSEATCKNVHFNNAHVTPHC 442
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 417 AITLNCQGVPCQGVIVQNIQL-----QNGR--AEQNNVQPAKGVVSPRC 460
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 8
 D96833
 hypothetical protein F18B13.25 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: D96833
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 anen, N.F.; Hughes, B.; Huizart, L.
 A:Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 kor, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: D96833
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-459 <STO>
 A:Cross-references: UNIPROT:O96862; GB:AEO05173; NID:G5902387; PIDN:AD5489.1; GSPDB:GN
 C:Gene: F18B3.25
 A:Map position: 1
 C:Superfamily: polygalacturonase
 Query Match 37.6%; Score 898.5; DB 2; Length 459;
 Best Local Similarity 40.4%; Pred. No. 1.6e-56;
 Matches 188; Conservative 82; Mismatches 146; Indels 49; Gaps 7;
 QY 9 LLLIIIFASSISTCRSNVIDNLFRQVYDNLGEFPAHDFQAYLSYLSKNIESNNIDKV 68
 DB LLLLLLVASSIALTAN-----ANSFSLIQLPRQSRSTR----- 46
 QY 69 DKNGIKVNVLSFGAKGDKTYVDNIAPQANEAASSSTPQVFPVPPKKNVLLQITFSG 128
 DB 47 -PRSERLHVNFQAKGVGVTDDTKAFADAKTKACSSKVKTRILVPEVYTCLLRPDLG 105
 QY 129 PCRSISVIGKISLEASS--KISDYKDRLLIAPDSVONLVVGGGCTINGNQWMPSSC 186
 DB 106 PCKARLTQISGTTIAPRDPDWEGLNKRKLYTHGLSRLVBEAGGTINGQGWEMERSC 165
 QY 187 KINKSLPCRDAPALTPWNCNKLKVNLSKNAQQIHIKFSCTNVVASNIMINASKSP 246
 DB 166 KHNHNSPCRGAPALTFFHKCKMREVNLVIDSQQMHIALTSCRRVITSLGLVIAPATSP 225
 QY 247 NTDGVHNSVNTQYIQISDITIGTDCISIVSGSQVQATNTTCRPHGISIGLSGNSSE 306
 DB 226 NTDGIIHISVSGIYDNTVSTGDCISIVKNSIQISLNIIIGPHGISIGLSGKSKSW 285
 QY 307 AYSNVTVNEMKIIIGAEVGRIKTWOGSSGOASNIKFLVEMOVQYPIIIDQNYCDRVE 366
 DB 286 EKVADIVDTALISDPANGVARIKTWOGSSGLVSKIIPIFNIGANNVNPITIIDQYCDSRK 345
 QY 367 PCIDQF-----SAVQKNVYENIKGTSATVAIKFDCSTNFPCEGIME 411
 DB 346 PCANQNLNTEPKLSTQKLTQTSAISINISFVHRTGSASKAKIKSCSDSPCRNILLQ 405
 QY 412 NINLVGSEGR-SEATCNVHFNNAEHV-THCTSLSEIDBALL 454
 DB 406 DIDLEPSWGDGFTESFCWEAYGSSGQVYPPPC-----LSDDTSFL 446
 RESULT 9
 H96728
 probable polygalacturonase F24UJ3.7 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: H96728
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federpiehl, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 anson, N.P.; Hughes, B.; Huizar, L.
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 kor, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H96728
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-468 <STO>
 A:Residues: 1-468 <STO>

```

C:Cross-references: UNIPROT:Q9CAL5; GB:AE005173; NID:96175142; PIDN:AA04869.1; GSPDB:GMK
A:Gene: F24013.7
A:Map position: 1

Query Match      35.2%; Score 840.5; DB 2; Length 468;
Best Local Similarity 41.6%; Pred. No. 2,3e-52;
Matches 177; Conservative 69; Mismatches 148; Indels 31; Gaps 7;

QY 37 DNIEQEAHDFQAVLS-YLSKNIESNNNIDKVDKNGIKVYNLSFGAKDGKTYDNIAP 95
DB 43 DDIEDDVTFPPDSFTSYSGKSL-----VNVDSFNASGDGVSDPTQAF 86
QY 96 EOANNEACSSRTFVQFVVPKKNLYLKQITFGSPCRSSISVKIFGSLSEASSKIS--DYKD 153
DB 87 IRAMTMACSAFNSV-LVLPQGRSYLVNATKFPGPCOEKILIQIDGTIIADDEPSQWDPKF 145
QY 154 RRLMIAFSPVQNLVYVGSGGTINGNGVWPPSSCKTIKSLPCRDAPALTFPNNCKNLVNN 213
DB 146 PRNLTQFSPKLGCVQNGNVIDIGSGTKMAALSKKNKNSPCVAPALTYSSNSVTVRG 205
QY 214 LKSKNAQOIHIKESCTNVVNASNLMINASAKSPNTDGVHVSNTQYIQISDTIIGTGDDCI 273
DB 206 LTIKRSQGMHLIQGSTYTRISRVWVTSPGDSPTNDGHIHTASTDVTVDQSKISTGDDCV 265
QY 274 SIVSGSNVQATNITTCGPHGISIGSLSGNSEAVYNSVTNEAKIIGAENGVRIRKTWOG 333
DB 266 SIVNSAKIKMKIRIYCGPHGISIGSLGQSHSGKGTVAVLLETAFLKNTTNGIRIKTWOG 325
QY 334 GSGGASNLKPLNVEMQVKKPIIIDQNYCDRVEPCIQSPSAVQKNVYENIKGTSATKV 393
DB 326 GNGYKVGKRFENVMQDVANPIIIDQFYCDSPSTCONQNSAVHISEIMYNNITGTVSSK 385
QY 394 AIKPEDSTNFPCEGIMENINLVGESGKPESEATCKNVHFNNAEH-----VTPHCTSLEIS 448
DB 386 AINFECSDAVPESHVILNNINIEGNDGK-VEAIC-----NSAAGFGGVVHPADCLYSH 439
QY 449 EDEAL 453
DB 440 DDKSL 444

RESULT 10
JC7100 polygalacturonase Cha o 2 - Japanese cypress
C:Species: Chamaecyparis obtusa (Japanese cypress)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: JC7100; FC7026
R:Motif, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
Biochem. Biophys. Res. Commun. 263, 166-177, 1999
A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major allel
A:Reference number: JC7100; MUID:99417540; PMID:10486272
A:Accession: JC7100
A:Molecule type: mRNA
A:Residues: 1-514 <MOR>
A:Cross-references: UNIPROT:Q7M1E7
A:Accession: FC7026
A:Molecule type: protein
A:Residues: 51-62 <MO2>

Query Match      33.3%; Score 796.5; DB 2; Length 514;
Best Local Similarity 38.7%; Pred. No. 3,5e-49;
Matches 172; Conservative 73; Mismatches 168; Indels 31; Gaps 7;

QY 8 ILLLIITASSISTGRSNVIDNLPKQVYDNILEGFADHFAVLYSLSKNISNNNDK 67
DB 12 LALQITLVMAA-----EDQSAQIMLDSIDIQ-----YL-----RSNRSLSKK 47
QY 68 V---DKNGIKVYNLVSFGAKDGKTYDNIAFEQANNEACSSRTFVQFVVPKKNLYLKQI 124
DB 48 LVHSHDADATYFNVQYGAVGDKHDSSTAFAITVMAACKKAASV-LVLPANKKFFVNNL 106
QY 125 TFSGCRSSISVYIKFGLSEASSKISDYKDRRLMIAFDSVQNLVVGGGGTINGNGQVWPPS 184

```

```

Db      107 VERGFCQPHLSFKVNDGTIAQDPDPAWKNKSKIMLOFAQITDPLNMGVGVIDGGQOQWMAQ 166
QY      185 SCK-INKSLPC--RDAPFALTFMNCNKLKVNNSKNAQOIHIFESCTNVVNASLMLINA 241
Db      167 QCKVNVGRVCDNRPRFPAIKIDVSKSTVAVKELTLMNSPEFHLVFGCEGKIGLKLKA 226
QY      242 SAKSPNTDGVYVSNTOYQISDTIIGTDDCISIVSGSQNVQANITCGHGISIGSLG 301
Db      227 PRDSPNTDGDIDIPASKRPHIEKVICGTDDCIAIGTSSNITTKIDLCGFGHGISIGSLG 286
QY      302 SGNSEAVYNTVNEAKTIGAEVGRITKTMQSGSGAASNTKFLNEMQDVKYPIIIDONY 361
Db      287 RDNRAEAVSHVVRNARFIDTQNGLRITKTMQSGSGSLASYITENVENINSENPILINQFY 346
QY      362 CDRVEPCIOQPSAVQVKNVYENIKGTSATKVAIKFPCSNFPCGEGIMENINLVGSGK 421
Db      347 CTSASACONQBSAVQIOGVYIKNHIGTSATPAALIQMCSDSVPTCTGIQLSNVLKLTSGK 406
QY      422 PSEATCKNVHFNNAEHVTPHCTSL 445
Db      407 PASCVDNARGFYSGRLIPTCKNL 430

```

RESULT 11

A:Accession: A96609
 Probable polygalacturonase P25P12.85 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 R:Accession: A96609
 R:Theologas, A.; Ecker, J.R.; Palm, C.D.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Ansen, N.F.; Hughes, B.; Hutzler, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L., Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A66141; MUID:21016719; PMID:11130712
 A:Accession: A96609
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-434 <STO>
 A:Cross-references: UNIPROT:Q9FXC1; GB:AE005173; NID:g9954742; PIDN:AAG09093.1; GSPDB:GN
 A:Gene: P25P12.85
 A:Map position: 1
 C:Superfamily: polygalacturonase

```

Query Match      32.9%; Score 787.5; DB 2; Length 434;
Best Local Similarity 41.3%; Pred. No. 1.2e-48;
Matches 170; Conservative 68; Mismatches 155; Indels 19; Gaps 8;

QY      54 YLSKNISNNNDIVDVKNGIKVINVLSFGAKGCKTYNDIAFEQANNEACSSRTT--VQF 111
Db      26 YLSRSPAPNAYNDNDIAFTVDFVTSFGAIGDCTDDTSAFQMAWDACMSTGKSMLL 85
QY      112 VPRNNKYYLKKQITFGSGCRSSISVKIFGSLA---SSKISDYKDRRMIMFSDVQVULV 167
Db      86 LVPTFCFLVKKPTTFNGPCRTNLVLDIDFIVSPDGRSPNSNY--QKQMMNFFYVNGLS 143
QY      168 VGGGGTINNGQVWMPSSCKINSL-----PCRDAPFALTFFMNCNKLKVNNLKSKNAQ 220
Db      144 IQSSGVINGQCKMMLPKCPHKGINGTQTGTGFC--DSPVAILRFPQSKVRLQGINPMNSA 202
QY      221 QIHKESECTNVVNASLMLINASAKSPNTDGVSNTOYQISDTIIGTDDCISIVSGSQ 280
Db      203 QEFVRFDCSDVYVSVIIKAPASSPNTDGIHIEHTNHNQIRNSNISGDDCISIGAGCF 262
QY      281 NVQANTITGCPGHGISIGSLSGNSEAVYNTVNEAKTIGAEVGRITKTMQSGSQASN 340

```

```

Db      263 NVDINVTCCGPHSGISIGSLGAEVNSQAYVSNITVTNSTINSDNGVRITKTMQSGSGSVR 322
QY      341 IKFLANEMQDVKYPIIIDONYCDREVEPCIOQPSAVQVKNVYENIKGT--SATKVAIKFPC 399
Db      333 IVFSNITLWVNRNIMIDQYCC--QTNNCANQTSAVIISDLVANIIGITDLSAPPIHFGC 381
QY      400 STNPFCEGIEMENINLVGSGKPSK--ATCGNVHFNNAEHVTPHCTSLSEISD 450
Db      382 SDSVPCINLITLVBVLDLFPSSQGHLENPFQWNAAGSMKTIIVPPYCLLDAPPD 433

```

RESULT 12

C:ry j II protein - Japanese cedar
 S:48730
 C:Species: Cryptomeria japonica (Japanese cedar)
 C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
 R:Accession: S48730
 R:Namba, M.; Kurose, M.; Torijoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kuri
 FEBS Lett. 353, 124-128, 1994
 A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar pc
 A:Reference number: S48730; MUID:95010777; PMID:7926035
 A:Accession: S48730
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1514 <NAM>
 A:Cross-references: UNIPROT:P43212; GB:D37765; NID:g577695; PIDN:BA07021.1; PID:d100759f

```

Query Match      32.9%; Score 787.5; DB 2; Length 514;
Best Local Similarity 38.5%; Pred. No. 1.6e-48;
Matches 173; Conservative 78; Mismatches 163; Indels 35; Gaps 9;

QY      8 ILLLIIPASSISTKRSNVVIDNLFKQYVDNILEQEFHDFQAVLSTLSKNISNNNIDK 67
Db      12 VMQQLIMAAA-----EDQSAQIMLSDIDQ-----YL-----RSNNSLRK 47
QY      68 VD---KNGIKYINVLSPGAKGDKTYNDIAFEQANNEACSSRTTPOFVPRKNYLLKQI 124
Db      48 VHSRHDALINIFVETKGAVDGDGHDCTEAFSTWQAAC--KKSAMLVFGNKKFPVNNVL 106
QY      125 TFSQGRSSISVKIFGSLAASSKISDYKDRRLMIAFSDVQVNLVVGGGTINNGQVWMP 184
Db      107 FFGPCQPHFTFKVDGIIAAVQNDPASWKNRIMLQFAKLITGLFMKGVIDDQCKQWMAQ 166
QY      185 SCK-INKSLPC--RDAPFALTFFMNCNKLKVNLSKNAQOIHIFESCTNVVNASLMLINA 241
Db      167 QCKVNVGRVCDNRPRFPAIKIDVSKSTVAVKELTLMNSPEFHLVFGNCGVYIKIGISITR 226
QY      242 SAKSPNTDGVYVSNTOYQISDTIIGTDDCISIVSGSQNVQATNITCGHGISIGSLG 301
Db      227 PRDSPNTDGDIDIPASKRPHIEKVICGTDDCIAIGTSSNITTKIDLCGFGHGISIGSLG 286
QY      302 SGNSEAVYNTVNEAKTIGAEVGRITKTMQSGSGAASNTKFLNEMQDVKYPIIIDONY 361
Db      287 RDNRAEAVSHVVRNARFIDTQNGLRITKTMQSGSGSLASYITENVENINSENPILINQFY 346
QY      362 CDRVEPCIOQPSAVQVKNVYENIKGTSATKVAIKFPCSNFPCGEGIMENINLVGSGK 421
Db      347 CTSASACONQBSAVQIOGVYIKNHIGTSATPAALIQMCSDSVPTCTGIQLSNVLKLTSGK 406
QY      422 PSEATCKNVHFNNAEHVTPHCTSL 448
Db      407 --IASCLNDNANGFYSGHVIIPACKNLSPS 433

```

RESULT 13

second major allergen Cry j II precursor - Japanese cedar
 J:JC2498
 C:Species: Cryptomeria japonica (Japanese cedar)
 C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
 R:Accession: JC2498; PC2346; A60147
 R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
 Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
 A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese c


```

Db      827 LEFSKLGKGVFOGKGVIDGSGSKWMAASCKKKS-----NALTESSGKVSGLTIQ 879
QY      218 NAOIHIKFESECTNVVNSMLMNASAKSPNTDGVHVSNTQYIOISPTIIGTGDDCISTVS 277
Db      880 NSQOMNFIARSDSVRVSKVWSSPGDSPNTDGIHITGSTNVLIDCKIGTDDCVSIVN 939
QY      278 GSONVQATNITCGPBGHGISIGSLGSGNSEAYVSNVTVNEAKIIGAENGVRIKTW----- 331
Db      940 ASSNIKMNKIYCGPBGHGISIGSLGKONTIGITQVLDLRLRETTNGRLIKTYOVHKK 999
QY      332 -----OQSGQASNIKEFLNEMQDVKYPPIIDONVCDRVE 366
Db      1000 KASLFSKNFPLARSTTLFHCFSFEQSGSGYOGIRFTVEMQDVANPILIDQFYCDSP 1059
QY      367 PCTIOQPSAVOVKNVVENIKGTSATKVAIKFDCSTNFPCEGIIMENINLVGESGKPSSEAT 426
Db      1060 TCONQTSAVAKISOIMYRNITGTTKSAKAIKFACSDPTVPCSHIVLNNVLENDGQ-VEAY 1118
QY      427 CKNVHFNNAEH-----VTPHCTSLSEDEAL 453
Db      1119 C-----NSAEGFGYGVHBSADCLYSHDKGL 1145

```

Search completed: March 24, 2005, 07:05:59
 Job time : 43 secs